GenCore version 4.5

OM protein - protein search, using sw model

Run on: September 13, 2002, 09:21:05 ; Search time 13.05 Seconds
(without alignments)
7.4[illegible] Million cell updates/sec

Title: ILE-LEU-ARG-GLY
Perfect score: 19
Sequence: 1 ILRG 4

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 231628 seqs, 24425594 residues

Total number of hits satisfying chosen parameters: 231628

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA.*
1: /cgn2_6/ptodata/2/ina/PA_COMB.pep *
2: /cgn2_6/ptodata/2/ina/PB_COMB.pep *
3: /cgn2_6/ptodata/2/ina/PA_COMB.pep *
4: /cgn2_6/ptodata/2/ina/PB_COMB.pep *
5: /cgn2_6/ptodata/2/ina/PCTUS_COMB.pep *
6: /cgn2_6/ptodata/2/ina/backfile1.pep *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	19	100.0	9	1	US-08-213-124-2	Sequence 27, Appl
2	19	100.0	9	2	US-08-482-651-18	Sequence 18, Appl
3	19	100.0	9	3	US-08-159-[illegible]	Sequence [illegible], App
4	19	100.0	9	4	US-08-660-092-2	Sequence 21, Appl
5	19	100.0	10	1	US-08-416-953-3	Sequence 3, Appli
6	19	100.0	10	2	US-08-469-[illegible]	Sequence [illegible], Appli
7	19	100.0	10	3	US-08-159-[illegible]	Sequence [illegible], App
8	19	100.0	12	2	US-08-591-444-3	Sequence 87, Appl
9	19	100.0	18	2	US-08-764-640-18	Sequence 140, App
10	19	100.0	18	3	US-08-804-449A-[illegible]	Sequence 107, App
11	19	100.0	18	3	US-08-973-225-2	Sequence 143, App
12	19	100.0	18	3	US-09-244-209A-[illegible]	Sequence 118, App
13	19	100.0	18	4	US-09-516-704-18	Sequence 140, App
14	19	100.0	19	1	US-08-672-671A-[illegible]	Sequence 11, Appl
15	19	100.0	68	4	US-09-025-151-[illegible]	Sequence 24, Appl
16	19	100.0	68	4	US-09-177-219-[illegible]	Sequence 158, App
17	19	100.0	70	4	US-09-367-958B-[illegible]	Sequence 29, Appl
18	19	100.0	70	4	US-09-367-958B-[illegible]	Sequence 85, Appl
19	19	100.0	70	4	US-09-367-958B-[illegible]	Sequence 88, Appl
20	19	100.0	70	4	US-09-367-958B-[illegible]	Sequence 89, Appl
21	19	100.0	70	4	US-09-467-958B-[illegible]	Sequence 90, Appl
22	19	100.0	71	1	US-08-280-441-[illegible]	Sequence [illegible], App
23	19	100.0	71	1	US-08-457-479-[illegible]	Sequence 13, Appl
24	19	100.0	71	1	US-08-555-628-[illegible]	Sequence 13, App
25	19	100.0	71	4	US-09-025-151-2	Sequence 25, Appl
26	19	100.0	71	5	PCT-US95-02795-[illegible]	Sequence 13, Appl
27	19	100.0	76	2	US-08-722-349-7	Sequence 7, Appl
28	19	100.0	76	4	US-09-204-528-7	Sequence 7, Appl
29	19	100.0	79	2	US-08-732-208-1	Sequence 1, Appl
30	19	100.0	79	4	US-09-425-[illegible]	Sequence [illegible], App
31	19	100.0	85	1	US-08-469-[illegible]	Sequence [illegible], App
32	19	100.0	95	1	US-08-250-975-7	Sequence 7, Appl
33	19	100.0	95	2	US-08-[illegible]	Sequence [illegible], App
34	19	100.0	95	2	US-08-[illegible]	Sequence [illegible], App
35	19	100.0	95	5	PCT-US94-10529	Sequence 7, Appl
36	19	100.0	112	3	US-08-345-8[illegible]	Sequence 11, App
37	19	100.0	117	4	US-08-545-8[illegible]	Sequence [illegible], App
38	19	100.0	119	4	US-09-470-473-1	Sequence [illegible], Appl
39	19	100.0	142	4	US-09-456-840-[illegible]	Sequence [illegible], Appl
40	19	100.0	142	4	US-09-456-840-1	Sequence [illegible], Appl
41	19	100.0	142	4	US-09-456-840-[illegible]	Sequence [illegible], Appl
42	19	100.0	150	2	US-08-851-188-1	Sequence 1, Appl
43	19	100.0	150	2	US-08-851-188-[illegible]	Sequence 1, Appl
44	19	100.0	151	4	US-09-255-[illegible]	Sequence [illegible], Appl
45	19	100.0	151	4	US-09-255-[illegible]	Sequence [illegible], Appl

ALIGNMENTS

RESULT 1
US-08-213-124-27
; Sequence 27, Application US/08213124
; Patent No. 5693325
; GENERAL INFORMATION:
; APPLICANT: Kahn, Michael
; TITLE OF INVENTION: Peptide Analogues and Libraries
; TITLE OF INVENTION: Thereof
; NUMBER OF SEQUENCES: 27
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: [illegible]
; STREET: [illegible]
; CITY: Seattle
; STATE: Washington
; COUNTRY: USA
; ZIP: 98104-7076
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #[illegible]
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/213,124
; FILING DATE: 15-MAR-1994
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Hermanns, Karl R.
; REGISTRATION NUMBER: 33,587
; REFERENCE/DOCKET NUMBER: 62-000,411
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 682-4900
; TELEFAX: (206) 682-6031
; TELEX: 4723836 SEED UW
; INFORMATION FOR SEQ ID NO: 27:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 9 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
US-08-213-124-27

Query Match [illegible]; Score 19; DB [illegible]; Length 9;
Best Local Similarity [illegible]; Pred. No. [illegible];
Matches [illegible]; Conservative [illegible]; Mismatches [illegible]; Indels [illegible]; Gaps [illegible];

Qy 1 ILRG 4
||||
Db 1 ILRG 4

```
;       FILING DATE:  06-JUN-1996
;       CLASSIFICATION:  445
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Park, Freddie K.
;       REGISTRATION NUMBER:  35,636
;       REFERENCE/DOCKET NUMBER:  25221 2060.1/3
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  (415) 813-5600
;       TELEFAX:  (415) 494-0792
;       TELEX:  706141
;   INFORMATION FOR SEQ ID NO:  21:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  9 amino acids
;       TYPE:  amino acid
;       STRANDEDNESS:  single
;       TOPOLOGY:  linear
;     IMMEDIATE SOURCE:
;       LIBRARY:  xyz (details pg. 36)
;       CLONE:  4E11
US-08-660-092-21


  Query Match             100.0%;  Score 19;  DB 5;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 1.7e+05;
  Matches    4;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 ILRG 4
           ||||
Db       4 ILRG 7


RESULT   5
US-08-416-950-3
; Sequence 3, Application US/08416950
; Patent No. 5780036
;  GENERAL INFORMATION:
;     APPLICANT:  CHISARI, Francis V.
;     TITLE OF INVENTION:  PEPTIDES FOR INDUCING [illegible]
;     TITLE OF INVENTION:  LYMPHOCYTIC RESPONSE  [illegible]
;     NUMBER OF SEQUENCES:  11
;     CORRESPONDENCE ADDRESS:
;       ADDRESSEE:  Townsend and Townsend Khourie and Crew
;       STREET:  One Market Plaza, Steuart Street Tower
;       CITY:  San Francisco
;       STATE:  CA
;       COUNTRY:  U.S.A.
;       ZIP:  94105-1492
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  [illegible]
;       SOFTWARE:  PatentIn Release #1.0, Version #1.30
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  [illegible]
;       FILING DATE:
;       CLASSIFICATION:  424
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  US
;       FILING DATE:
;       APPLICATION NUMBER:  US 07/945,808
;       FILING DATE:  26 AUG 1992
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  US 07/749,540
;       FILING DATE:  26 AUG 1991
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Parmelee, Steven W.
;       REGISTRATION NUMBER:  31,990
;       REFERENCE/DOCKET NUMBER:  11740-2-2
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  (206) 467-9600
;       TELEFAX:  (415) 543-5043
;   INFORMATION FOR SEQ ID NO:  3:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  10 amino acids
;       TYPE:  amino acid
;       STRANDEDNESS:  single
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  peptide
US-08-416-950-3


  Query Match             [illegible]%;  Score 19;  DB 5;  Length 10;
  Best Local Similarity   [illegible]%;  Pred. No.  3;
  Matches    4;  Conservative    [illegible];  Mismatches    [illegible];  Indels    [illegible];  Gaps    [illegible];

Qy       1 ILRG 4
           |||.|
Db       1 ILRG 4


RESULT   6
US-08-469-830-3
; Sequence 3, Application US/08469830
; Patent No. 5942224
;  GENERAL INFORMATION:
;     APPLICANT:  CHISARI, Francis V.
;     TITLE OF INVENTION:  [illegible] INDUCING CYTOTOXIC
;     TITLE OF INVENTION:  LYMPHOCYTIC RESPONSE [illegible] HEPATITIS VIRUS
;     NUMBER OF SEQUENCES:  11
;     CORRESPONDENCE ADDRESS:
;       ADDRESSEE:  [illegible]
;       STREET:  One Market Plaza, Steuart Street
;       CITY:  San Francisco
;       STATE:  CA
;       COUNTRY:  U.S.A.
;       ZIP:  94105-1492
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  PatentIn Release #1.0, Version #1.30
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/[illegible]
;       FILING DATE:  [illegible]
;       CLASSIFICATION:  424
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  US 08/[illegible]
;       FILING DATE:  2 AUG 1995
;       APPLICATION NUMBER:  US 07/[illegible]
;       FILING DATE:  26 AUG 1992
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  US 07/749,540
;       FILING DATE:  26 AUG 1991
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Parmelee, Steven W.
;       REGISTRATION NUMBER:  31,990
;       REFERENCE/DOCKET NUMBER:  11740-2-2
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  (206) 467-9600
;       TELEFAX:  (415) 543-5043
;   INFORMATION FOR SEQ ID NO:  3:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  10 amino acids
;       TYPE:  amino acid
;       STRANDEDNESS:  single
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  peptide
US-08-469-830-3


  Query Match             [illegible]%;  Score 19;  DB 5;  Length 10;
  Best Local Similarity   [illegible]%;  Pred. No.  3;
  Matches    4;  Conservative    [illegible];  Mismatches    [illegible];  Indels    [illegible];  Gaps    [illegible];

Qy       1 ILRG 4
           |||.
```

```
;       MEDIUM TYPE:  Floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  PatentIn Release #1.0, Version #1.30
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/764,640
;       FILING DATE:  11-DEC-1996
;       CLASSIFICATION:  514
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Hrubiec, Robert J.
;       REGISTRATION NUMBER:  36,492
;       REFERENCE/DOCKET NUMBER:  PK4281
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  919-248-1000
;   INFORMATION FOR SEQ ID NO:  140:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  18 amino acids
;       TYPE:  amino acid
;       STRANDEDNESS:
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  peptide
US-08-764-640-140


  Query Match             100.0%;  Score 19;  DB 11;  Length 14;
  Best Local Similarity   100.0%;  Pred. No. 61;
  Matches    4;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 ILEG 4
           ||||
Db      11 ILEG 14


RESULT 10
US-08-804-439A-107
; Sequence 107, Application US/08/804,439A
; Patent No. 6015565
;  GENERAL INFORMATION:
;     APPLICANT:  Rose, Timothy M.
;     APPLICANT:  Bosch, Marnix L.
;     APPLICANT:  Strand, Kurt
;     TITLE OF INVENTION:  GLYCOPROTEIN B OF THE EQUINE
;     TITLE OF INVENTION:  SUBFAMILY OF HERPES VIRUSES
;     NUMBER OF SEQUENCES:  113
;     CORRESPONDENCE ADDRESS:
;       ADDRESSEE:  Fish & Richardson P.C.
;       STREET:  4225 Executive Square, Suite
;       CITY:  La Jolla
;       STATE:  CA
;       COUNTRY:  USA
;       ZIP:  92037
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  PatentIn Release #1.0, Version #1.30
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/804,439A
;       FILING DATE:  February 21, 1997
;       CLASSIFICATION:  424
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Halle, Lisa A.
;       REGISTRATION NUMBER:  38,347
;       REFERENCE/DOCKET NUMBER:  06378/200400
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  (619) 678-5070
;       TELEFAX:  (619) 678-5099
;       TELEX:
;   INFORMATION FOR SEQ ID NO:  107:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  18 amino acids
;       TYPE:  amino acid
;       STRANDEDNESS:  single
;       TOPOLOGY:  linear
US-08-804-439A-107


  Query Match             100.0%;  Score 19;  DB 11;  Length 14;
  Best Local Similarity   100.0%;  Pred. No. 61;
  Matches    4;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 ILEG 4
           ||||
Db       5 ILEG 8


RESULT 11
US-08-974-225-140
; Sequence 140, Application US/08974225A
; Patent No. 6086876
;   GENERAL INFORMATION:
;       APPLICANT: Esser, Wolfgang
;                  Barrett, Ronald W.
;                  Saxe, Steven
;                  Dillon, David C.
;                  Gates, Christian
;                  Basorcdon, Sherrill S.
;                  Matthews, Larry C.
;                  Schatz, Peter J.
;                  Wagstrom, Christopher E
;                  Wrighton, N. Thomas
;       TITLE OF INVENTION: PEPTIDES AND COMPOUNDS THAT BIN     A
;                           IBP MOLECULES
;       NUMBER OF SEQUENCES: 174
;       CORRESPONDENCE ADDRESS:
;            ADDRESSEE: Lyon & Lyon
;            STREET: 633 West Fifth Street, Suite
;            CITY: Los Angeles
;            STATE: CA
;            COUNTRY: USA
;            ZIP: 90071
;       COMPUTER READABLE FORM:
;            MEDIUM TYPE: Floppy disk
;            COMPUTER: IBM PC compatible
;            OPERATING SYSTEM: PC-DOS/MS-DOS
;            SOFTWARE: PatentIn Release #1.0, Version #1.30
;       CURRENT APPLICATION DATA:
;            APPLICATION NUMBER: US/08/974,225
;            FILING DATE: 04-Dec-1997
;       ATTORNEY/AGENT INFORMATION:
;            NAME: Hrubiec, Robert J.
;            REGISTRATION NUMBER: 36,492
;            REFERENCE/DOCKET NUMBER: 183/05 US
;       TELECOMMUNICATION INFORMATION:
;            TELEPHONE: 619-210-1900
;   INFORMATION FOR SEQ ID NO: 140:
;       SEQUENCE CHARACTERISTICS:
;            LENGTH: 18 amino acids
;            TYPE: amino acid
;            STRANDEDNESS: Unknown
;            TOPOLOGY: Linear
;       MOLECULE TYPE: peptide
;       SEQUENCE DESCRIPTION: Seq ID NO: 140:
US-08-974-225-140


  Query Match             100.0%;  Score 19;  DB 11;  Length 14;
  Best Local Similarity   100.0%;  Pred. No. 61;
  Matches    4;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 ILEG 4
           ||||
Db      11 ILEG 14


RESULT 12
```

```
;       FILING DATE:  28 JUNE 1996
;       CLASSIFICATION:  435
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  WEINER, Marc S.
;       REGISTRATION NUMBER:  42,181
;       REFERENCE/DOCKET NUMBER:  1422-02641
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  (703) 205 8000
;       TELEFAX:  (703) 205-8050
;       TELEX:  248345
;   INFORMATION FOR SEQ ID NO:  11:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  19 amino acids
;       TYPE:  amino acid
;       STRANDEDNESS:  single
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  peptide
;     FRAGMENT TYPE:  internal fragment
US-08-672-571A-11


  Query Match             100.0%;  Score 19;  DB 4;  Length 19;
  Best Local Similarity   100.0%;  Pred. No. 7.4;
  Matches    4;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 ILRG 4
           ||||
Db       8 ILRG 11


RESULT  15
US-09-025-151-24
; Sequence 24, Application US/09025151
; Patent No. 6187535
; GENERAL INFORMATION:
;  APPLICANT: Legrain, Pierre
;  APPLICANT: Fromont, Micheline
;  APPLICANT: Rain, Jean-Christophe
;  TITLE OF INVENTION: FAST AND EXHAUSTIVE METHOD FOR SELECTING A PREY
;  TITLE OF INVENTION: POLYPEPTIDE INTERACTING WITH A BAIT POLYPEPTIDE OF
;  TITLE OF INVENTION: INTEREST: APPLICATION TO THE CONSTRUCTION OF MAPS OF
;  TITLE OF INVENTION: INTERACTORS POLYPEPTIDES
;  FILE REFERENCE: 03495-0164
;  CURRENT APPLICATION NUMBER: US/09/025,151
;  CURRENT FILING DATE: 1998-02-18
;  NUMBER OF SEQ ID NOS: 29
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 24
;   LENGTH: 68
;   TYPE: PRT
;   ORGANISM: Saccharomyces cerevisiae
US-09-025-151-24


  Query Match             100.0%;  Score 19;  DB 3;  Length 68;
  Best Local Similarity   100.0%;  Pred. No. 2.3e+02;
  Matches    4;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 ILRG 4
           ||||
Db      22 ILRG 25



Search completed: September 13, 2002, 09:41:59
Job time: 644 sec
```

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

SM protein - protein search, using sw model

Run on: September 13, 2002, 09:56:[illegible] ; [illegible] Seconds
(without alignments)
27.610 Million cell updates/sec

Title: ILE-LEU-ARG-GLY
Perfect score: 19
Sequence: 1 ILRG 4

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 562222 seqs, 172994529 residues

Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_19:*
1: sp_archea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_rvirus:*
16: sp_bacteriap:*
17: sp_archeap:*

Pred. No. is the number of results predicted to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distributions.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	19	100.0	23	6	Q9TRI4	Q9TRI4 bos taurus
2	19	100.0	28	15	O71995	O71995 human endog
3	19	100.0	35	4	Q96EZ3	Q96EZ3 homo sapien
4	19	100.0	36	15	Q76042	Q76042 human immun
5	19	100.0	37	13	P82839	P82839 rana berlan
6	19	100.0	39	6	Q9MZF5	Q9MZF5 bos taurus
7	19	100.0	47	10	Q93YL3	Q93YL3 arabidopsis
8	19	100.0	49	5	Q26787	Q26787 trypanosoma
9	19	100.0	50	1	O93677	O93677 sulfolobus
10	19	100.0	52	5	Q26173	Q26173 plasmodium
11	19	100.0	53	2	Q9RDF8	Q9RDF8 streptomyce
12	19	100.0	53	16	Q9KCE5	Q9KCE5 bacillus ha
13	19	100.0	56	7	Q9MW66	Q9MW66 alcolapia a
14	19	100.0	58	5	Q18191	Q18191 caenorhabd
15	19	100.0	58	5	Q26174	Q26174 plasmodium
16	19	100.0	60	11	Q61424	Q61424 mus musculu
17	19	100.0	[illegible]	[illegible]	[illegible]	[illegible]
18	19	100.0	[illegible]	[illegible]	Q9KII5	[illegible]
19	19	100.0	[illegible]	[illegible]	Q9KVK9	[illegible]
20	19	100.0	[illegible]	[illegible]	Q9AEQ2	[illegible]
21	19	100.0	[illegible]	[illegible]	[illegible]	[illegible]
22	19	100.0	[illegible]	[illegible]	[illegible]	[illegible]
23	19	100.0	[illegible]	[illegible]	[illegible]	[illegible]
24	19	100.0	[illegible]	[illegible]	Q9TLY2	[illegible]
25	19	100.0	[illegible]	[illegible]	Q967S6	[illegible]
26	19	100.0	[illegible]	[illegible]	Q9PIH2	[illegible]
27	19	100.0	[illegible]	[illegible]	Q99II9	[illegible]
28	19	100.0	[illegible]	[illegible]	Q9VXE1	[illegible]
29	19	100.0	[illegible]	[illegible]	Q9WEI4	[illegible]
30	19	100.0	[illegible]	[illegible]	Q9NGC4	[illegible]
31	19	100.0	[illegible]	[illegible]	[illegible]	[illegible]
32	19	100.0	[illegible]	[illegible]	Q96JZ5	[illegible]
33	19	100.0	[illegible]	[illegible]	[illegible]	[illegible]
34	19	100.0	[illegible]	[illegible]	[illegible]	[illegible]
35	19	100.0	[illegible]	[illegible]	[illegible]	[illegible]
36	19	100.0	[illegible]	[illegible]	[illegible]	[illegible]
37	19	100.0	[illegible]	[illegible]	[illegible]	[illegible]
38	19	100.0	[illegible]	[illegible]	Q9XKI4	[illegible]
39	19	100.0	[illegible]	[illegible]	[illegible]	[illegible]
40	19	100.0	[illegible]	[illegible]	[illegible]	[illegible]
41	19	100.0	[illegible]	[illegible]	[illegible]	[illegible]
42	19	100.0	[illegible]	[illegible]	[illegible]	[illegible]
43	19	100.0	[illegible]	[illegible]	[illegible]	[illegible]
44	19	100.0	[illegible]	[illegible]	42755	[illegible]
45	19	100.0	[illegible]	[illegible]	42752	[illegible]

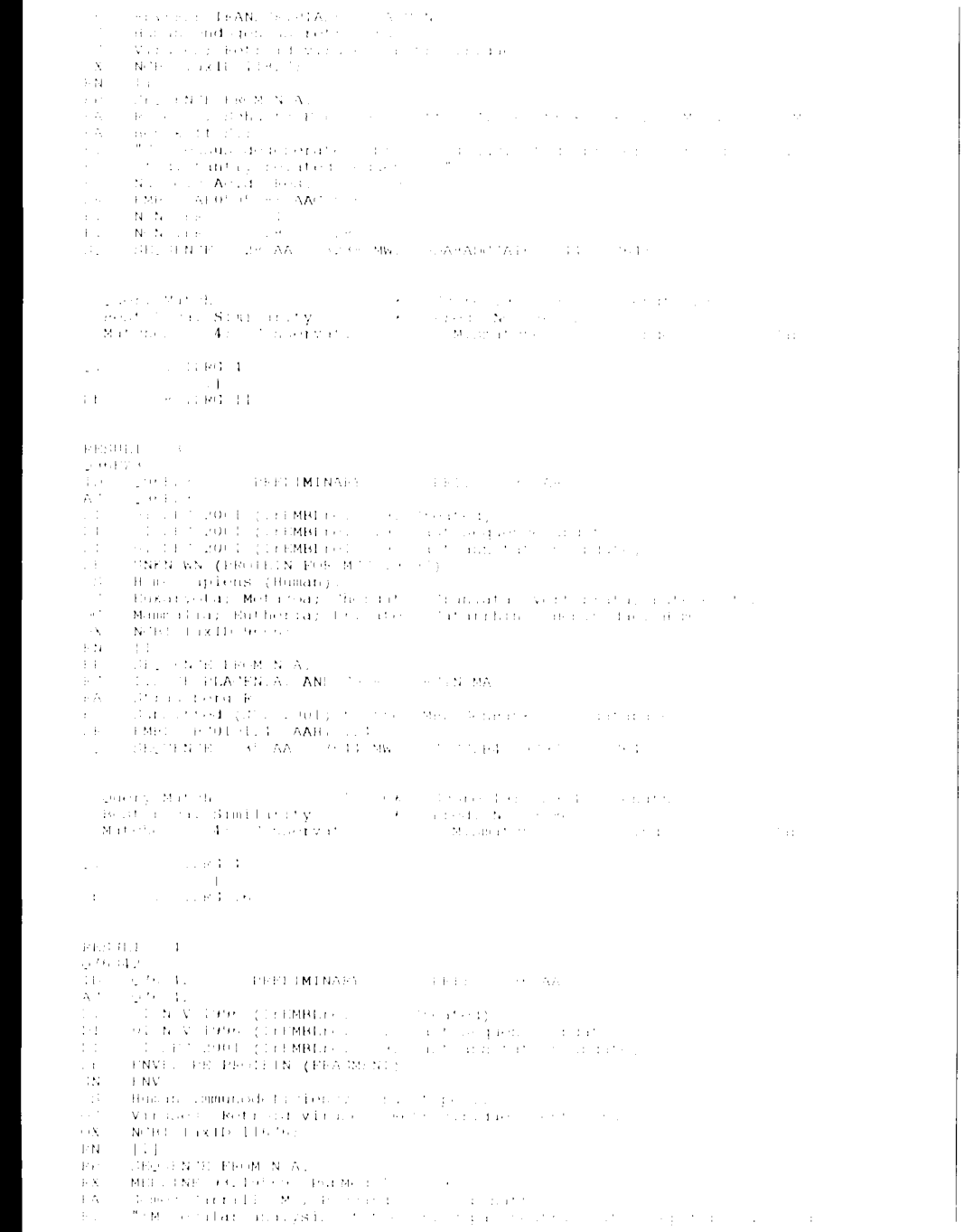
ALIGNMENTS

RESULT 1
Q9TRI4
ID Q9TRI4 [illegible] [illegible]
AC Q9TRI4;
DT 01-MAY 2000 (TrEMBLrel. 13, Created)
DT 01-MAY 2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN 2000 (TrEMBLrel. 14, Last annotation update)
DE AMINOPEPTIDASE N (EPAMINE).
OS Bos taurus (bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; [illegible]
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; [illegible]
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE.
RX MEDLINE=97157143; PubMed=8990315;
RA Plakidou-Dymock S., Dymock D., Hooper N.M.;
RT "A role for aminopeptidase N in peptide [illegible]";
RT in bovine renal brush border membranes.";
RL Biochem. J. 330:59-65(1998).
SQ SEQUENCE 23 AA; 2701 MW; [illegible];

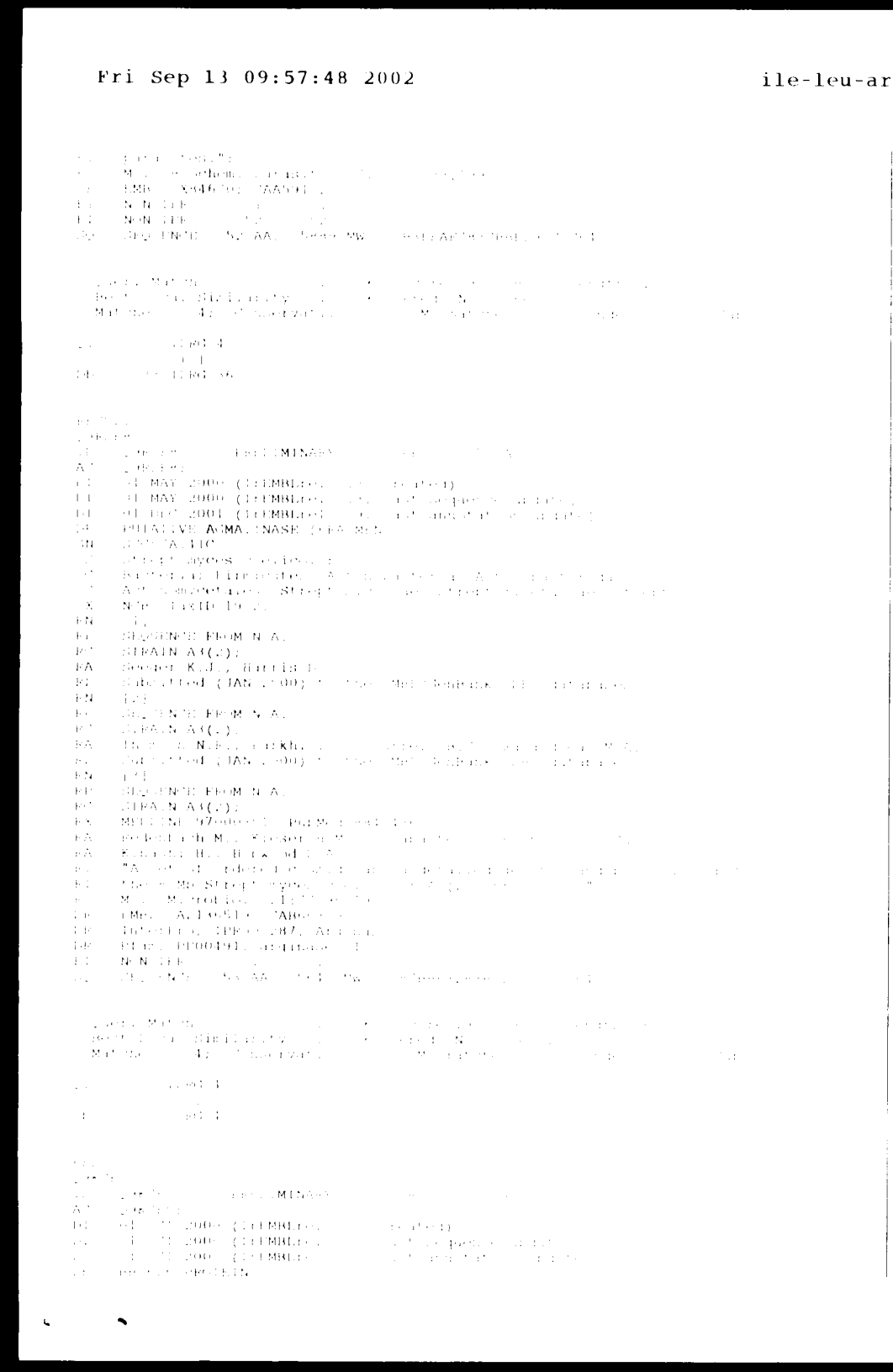
Query Match 100.0%; Score 19; [illegible];
Best Local Similarity 100.0%; Pred. No. [illegible];
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

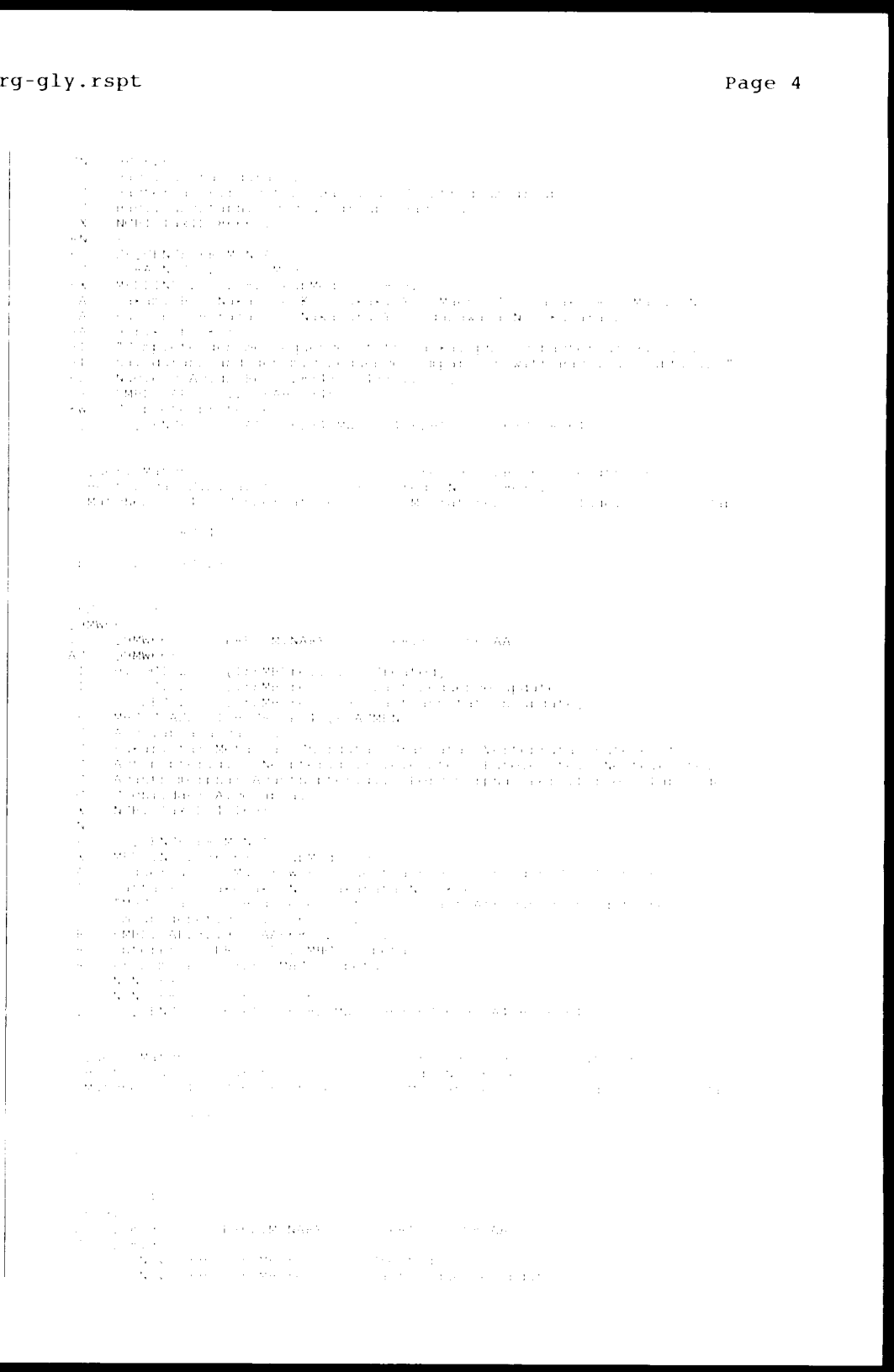
Qy 1 ILRG 4
||||
Db 16 ILRG 19

RESULT 2
O71995
ID O71995 [illegible]
AC O71995;
DT 01-AUG 1998 (TrEMBLrel. 7, Created)
DT 01-AUG 1998 (TrEMBLrel. 7, Last sequence update)
DT 01-DEC 2001 (TrEMBLrel. 19, Last annotation update)



```
OC   Mammalia; Eutheria; Cetartiodactyla; Ruminan' ... Pecora; Bovidae;
OC   Bovidae; Bovinae; Bos.
OX   NCBI_TaxID=9913;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE 20280829; PubMed 10920210;
RA   Sonstegard T.S., Garrett W.M., Ashwell M.S., ... G.L.,
RA   Kappes S.M., Van Tassell C.P.;
RT   "Comparative map alignment of BTA27 and HSA1 ... identify
RT   conserved segments of genome containing fat ... position QTL.";
RL   Mamm. Genome 11:682-688(2000).
DR   EMBL; AF230197; AAF79127.1; -.
DR   BSSP; P07688; 1qbg.
DR   InterPro; IPR000668; Peptidase_C1.
DR   Pfam; PF00112; Peptidase_C1; 1.
FT   NON_TER       1       1
SQ   SEQUENCE   49 AA;  4385 MW;  B4B2514E50F014F... CG43

  Query Match           100.0%;  Score 19;  ... 4;  ... 4;  E(3473) 4%
  Best Local Similarity  100.0%;  Pred. No. 1.2e-...
  Matches    4;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0

Qy       1 ILRG 4
           ||||
Db      16 ILRG 19


RESULT   7
Q93Y13
ID   Q93Y13      PRELIMINARY;      PRT;    47 AA.
AC   Q93Y13;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation ... update)
DE   L-ASPARTATE OXIDASE-LIKE PROTEIN.
GN   AT5G14760 OR T9L3_60.
OS   Arabidopsis thaliana (Mouse-ear cress).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; ... core eudicots; Rosidae;
OC   eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX   NCBI_TaxID=3702;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Nagych M., Karlin-Neumann G., Southwick A., ... R., Miranda M.,
RA   Palm C.J., Bowser L., Jones T., Hanh J., ... P., Chen H.,
RA   Cheuk R., Chung M.K., Hayeshizaki Y., Ishiba ... Kamiya A., Kawai J.,
RA   Kim C., Lin J., Liu S.X., Narusaka M., Pham ... Fukuda H.,
RA   Sakurai T., Sato M., Seki M., Shinn P., ... Shinozaki K.,
RA   Ecker J., Theologis A., Davis R.W.;
RL   Submitted (OCT-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AY059800; AAL24282.1; -.
SQ   SEQUENCE   47 AA;  5415 MW;  5057AC7D77CE6... CG43

  Query Match           100.0%;  Score 19;  ... 4;  ... 4;
  Best Local Similarity  100.0%;  Pred. No. 5.1e-...
  Matches    4;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0

Qy       1 ILRG 4
           ||||
Db      11 ILRG 14


RESULT   6
Q26787
ID   Q26787      PRELIMINARY;      PRT;    49 AA.
AC   Q26787;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation ... update)
DE   PAG3 PROTEIN.
GN   PAG3.
OS   Trypanosoma brucei.
OC   Eukaryota; Euglenozoa; Kinetoplastida; ... Trypanosomatidae; Trypanosoma.
OX   NCBI_TaxID=5691;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN 427;
RA   Vassella E., Yamage M., Krisinger J.;
RL   Submitted (OCT-1994) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; L26656; AAA30120.1; -.
SQ   SEQUENCE   49 AA;  5115 MW;  1E9CC1D6C8... CG43

  Query Match           ... 4;  Score ...;  ... 4;  ... 4;
  Best Local Similarity  ... 4;  Pred. No. ...
  Matches    4;  Conservative    ...;  Mismatches    ...;  Indels    ...;  Gaps ...

Qy       1 ILRG 4
           ||||
Db      10 ILRG 13


RESULT   9
Q93677
ID   Q93677      PRELIMINARY;      PRT;    ... AA.
AC   Q93677;
DT   01-MAY-1999 (TrEMBLrel. ..., Created)
DT   01-MAY-1999 (TrEMBLrel. ..., Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. ..., Last annotation update)
DE   HYPOTHETICAL ...7 KDA PROTEIN.
OS   Sulfolobus sp. NOB8H.
OG   Plasmid pNOB8.
OC   Archaea; Crenarchaeota; ... Desulfurococcales; ... Sulfolobus.
OX   NCBI_TaxID=81600;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=NOB8H2;
RX   MEDLINE 99041580; PubMed ...;
RA   She Q., Phan H., Garrett R.A., Albers S.-V., Stedman K.M., ...;
RT   "Genetic profile of pNOB8 from Sulfolobus ... first ... plasmid ...";
RT   plasmid from an archaeon.";
RL   Extremophiles 2:417-425(1998).
DR   EMBL; AJ010435; CAA...;
KW   Hypothetical protein; Plasmid.
SQ   SEQUENCE   ... AA;  ... MW;  ... CRC64;

  Query Match           ... 4;  Score ...;  ... 4;  ...
  Best Local Similarity  ... 4;  Pred. No. ...
  Matches    4;  Conservative    ...;  Mismatches    ...;  Indels    ...;  Gaps ...

Qy       1 ILRG 4
           |:|
Db      44 ILRG 47


RESULT   10
Q26173
ID   Q26173      PRELIMINARY;      PRT;    ... AA.
AC   Q26173;
DT   01-NOV-1996 (TrEMBLrel. ..., Created)
DT   01-NOV-1996 (TrEMBLrel. ..., Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. ..., Last annotation update)
DE   GAM1 PROTEIN (PARMEN).
GN   GAM1.
OS   Plasmodium vivax.
OC   Eukaryota; Alveolata; Apicomplexa; Haemosporina; Plasmodiidae;
OX   NCBI_TaxID=5855;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE 96358750; PubMed ...;
RA   Snewin V.A., Khouri E., Wattavidanage ...;
RA   Mendis K.N., Carter ...;
RT   "A new polymorphic ... from ... vivax ...
```





```
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   HYPOTHETICAL 7.1 KDA PROTEIN.
GN   C26B2.5.
OS   Caenorhabditis elegans.
OC   Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC   Rhabditidae; Peloderinae; Caenorhabditis.
OX   NCBI_TaxID=6239;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BRISTOL N2;
RX   MEDLINE=99069614; PubMed=9851916;
RA   None;
RT   "Genome sequence of the nematode C. elegans: a platform for
RT   investigating biology. The C. elegans Sequencing Consortium.";
RL   Science 282:2012-2018(1998).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN BRISTOL N2;
RA   Du Z., Cattuno S.;
RT   "The sequence of C. elegans cosmid C26B2.";
RL   Submitted (DEC-1995) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE FROM N.A.
RC   STRAIN BRISTOL N2;
RA   Waterston R.;
RT   "Direct Submission.";
RL   Submitted (AUG-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; Z41554; AAC24259.1; -.
KW   Hypothetical protein.
SQ   SEQUENCE   58 AA;  6088 MW;  8B1DF095B516A1C4 CRC64;
```

```
  Query Match             100.0%;  Score 19;  DB 1;  Length 58;
  Best Local Similarity   100.0%;  Pred. No. 6.5e+03;
  Matches    4;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 ILRG 4
           ||||
Db       5 ILRG 8
```

```
RESULT  15
Q26174
ID   Q26174       PRELIMINARY;      PRT;    58 AA.
AC   Q26174;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   GAM1 PROTEIN (FRAGMENT).
GN   GAM1.
OS   Plasmodium vivax.
OC   Eukaryota; Alveolata; Apicomplexa; Haemosporina; Plasmodium.
OX   NCBI_TaxID=5855;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=95356700; PubMed=7644477;
RA   Snewin V.A., Khouri E., Wattavidanage J., Perera L., Premawansa S.,
RA   Mendis K.N., David P.H.;
RT   "A new polymorphic marker for PCR typing of Plasmodium vivax
RT   parasites.";
RL   Mol. Biochem. Parasitol. 71:135-138(1995).
DR   EMBL; X84674; CAA59195.1; -.
FT   NON_TER       1       1
FT   NON_TER      58      58
SQ   SEQUENCE   58 AA;  6411 MW;  E3BB144F1E7AC1E8 CRC64;
```

```
  Query Match             100.0%;  Score 19;  DB 1;  Length 58;
  Best Local Similarity   100.0%;  Pred. No. 6.5e+03;
  Matches    4;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 ILRG 4
           ||||
```

```
Db      49 ILRG 42


Search completed: September 13, 2002, 09:57:42
Job time: 114 secs
```

GenCore version 4.5

OM protein - protein search, using sw model

Run on: September 13, 2002, 09:29:45 ; Search time 10.33 Seconds
(without alignments)
14.66 Million cell updates/sec

Title: ILE-LEU-ARG-GLY
Perfect score 19
Sequence: 1 ILRG 4

Scoring table BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 105224 seqs, 38719650 residues

Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 5%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	19	100.0	34	1	LPIN_PROVU	P28779 proteus vul
2	19	100.0	43	1	BDS1_ANESU	P11494 anemonia su
3	19	100.0	52	1	YP55_ECOLI	P14505 escherichia
4	19	100.0	64	1	OCML_ECOLI	P26770 escherichia
5	19	100.0	75	1	YF62_THEMA	Q9X1Q3 thermotoga
6	19	100.0	76	1	PGXC_HUMAN	Q15847 homo sapien
7	19	100.0	77	2	PGXC_YEAST	P40204 saccharomyc
8	19	100.0	90	2	OC2_PETHY	P43296 petunia hyb
9	19	100.0	94	1	YC76_ARCFU	O28992 archaeoglob
10	19	100.0	99	1	ITI2_SINAL	P26780 sinapis alb
11	19	100.0	101	1	VPI5_BPAPS	Q91L13 bacteriopha
12	19	100.0	105	1	UL03_HCMVA	P16775 human cytom
13	19	100.0	107	1	VNS1_PVM	P28388 pneumonia v
14	19	100.0	116	1	YJG7_YEAST	P40365 saccharomyc
15	19	100.0	117	1	RL24_METTE	O26122 methanobact
16	19	100.0	122	1	YH11_AZOBR	P26724 azospirillu
17	19	100.0	125	1	MERR_STRLI	P35546 streptomyce
18	19	100.0	129	1	NKL_PIG	Q29075 sus scrofa
19	19	100.0	130	1	UL33_HSV11	P13217 herpes simp
20	19	100.0	142	1	RR13_PORPU	P51292 porphyra pu
21	19	100.0	144	1	RL13_AQUAE	O67722 aquifex aeo
22	19	100.0	145	1	MCPL_SYNY3	P73553 synechocyst
23	19	100.0	145	1	RL13_BACSU	P76074 bacillus su
24	19	100.0	145	1	RL13_STACA	Q8NX90 staphylococ
25	19	100.0	146	1	RL13_BORBU	O51514 borrelia bu
26	19	100.0	147	1	IL4_RAT	P20798 rattus norv
27	19	100.0	147	1	RL13_STRCO	Q58874 streptomyce
28	19	100.0	147	1	YAU8_SCHPO	Q10163 schizosacch
29	19	100.0	148	1	MBBI_ECOLI	P08097 escherichia
30	19	100.0	148	1	NRD1_CORGL	Q9XD64 corynebacte
31	19	100.0	149	1	RL13_CHLPN	Q9Z8T7 chlamydia p
32	19	100.0	150	1	CNRD_BOVIN	Q95142 bos taurus
33	19	100.0	150	1	CNRE_HUMAN	Q43764 homo sapien
34	19	100.0	150	1	CNRE_MOUSE	[illegible]
35	19	100.0	150	1	RL13_CHLMU	[illegible]
36	19	100.0	153	1	RL13_CHLTR	[illegible]
37	19	100.0	154	1	RL13_SYNY3	[illegible]
38	19	100.0	154	1	YV46_BACHD	[illegible]
39	19	100.0	157	1	PSAE_SYNY3	[illegible]
40	19	100.0	157	1	YP09_YEAST	[illegible]
41	19	100.0	161	1	Y_AD_BACSU	[illegible]
42	19	100.0	166	1	Y_AD_SYNY3	[illegible]
43	19	100.0	172	1	YE74_NEIGO	[illegible]
44	19	100.0	177	1	PPC_PANEA	[illegible]
45	19	100.0	178	1	RL13_SULSO	[illegible]

ALIGNMENTS

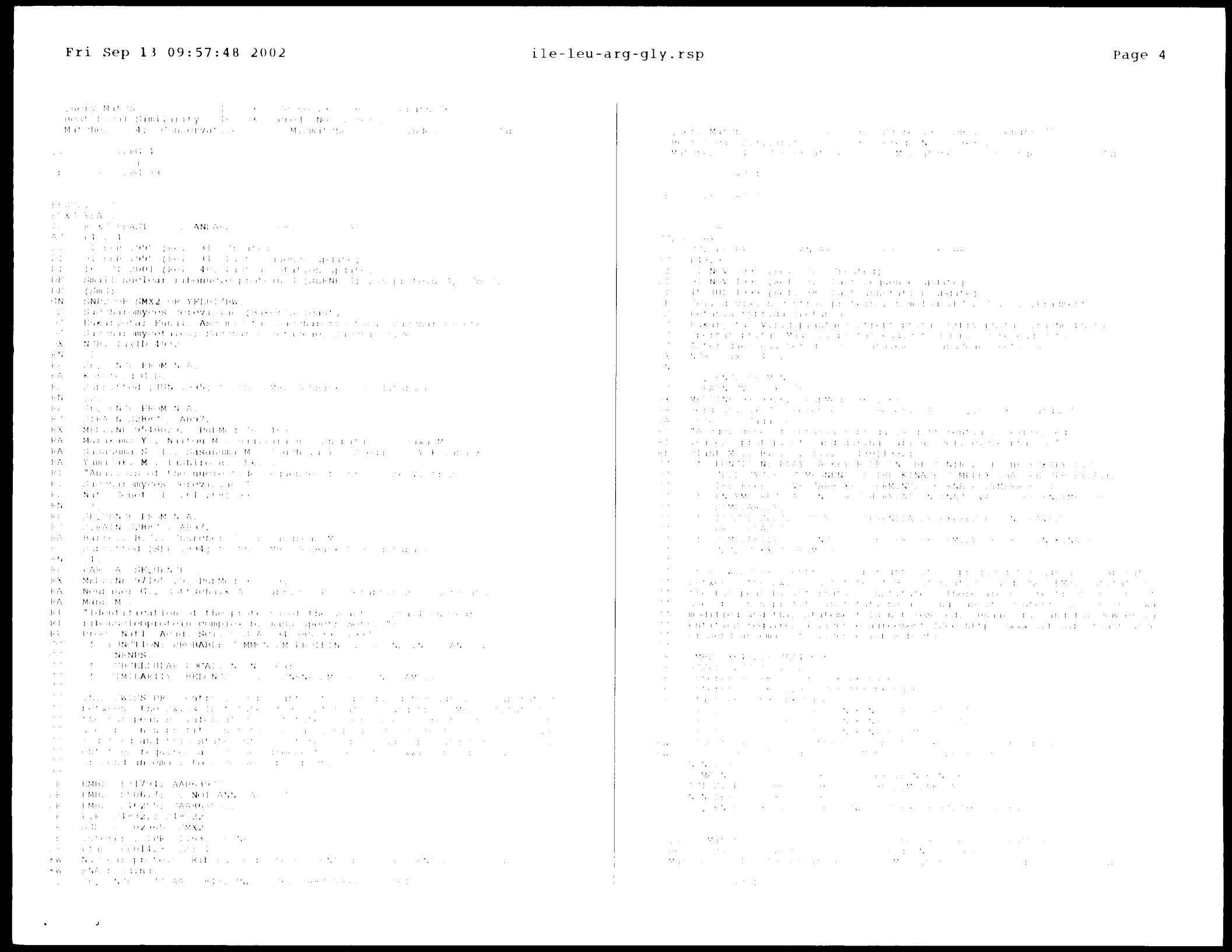
RESULT 1
LPIN_PROVU
ID LPIN_PROVU STANDARD; PRT; 34 AA.
AC P28779;
DT 01-DEC-1992 (Rel. 25, Created)
DT 01-DEC-1992 (Rel. 25, Last sequence update)
DT 01-OCT-1993 (Rel. 27, Last annotation update)
DE Tryptophanase leader peptide.
GN INAL OR INAT.
OS Proteus vulgaris.
OC Bacteria; Proteobacteria; Gamma subdivision; Enterobacteriales;
OC Proteus.
OX NCBI_TaxID=585;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE 93019056; PubMed=1400414;
RA Kamath A.V., Yanofsky C.;
RT "Characterization of the tryptophanase operon leader region and
RT cloning, nucleotide sequence, and analysis of the tnaL gene.";
RT synthesis of the leader peptide and termination of transcription.";
RL J. Biol. Chem. 267:22669-22676(1992).
CC -!- [illegible]
CC -!- This SWISS-PROT entry is copyright. It is [illegible] terms
CC between the Swiss Institute of Bioinformatics and the [illegible]
CC the European Bioinformatics Institute. There are [illegible]
CC use by non-profit [illegible] as long as [illegible]
CC modified and this statement [illegible]
CC entities requires a [illegible]
CC or send an email to [illegible]
CC [illegible]
DR EMBL; M93477; AAA[illegible]
DR PIR; A44038; A44038.
KW Leader peptide.
SQ SEQUENCE 34 AA; [illegible] MW; [illegible] CRC64;

Query Match [illegible]; Score 19; DB 1; Length 34;
Best Local Similarity [illegible]; Pred. No. [illegible]
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

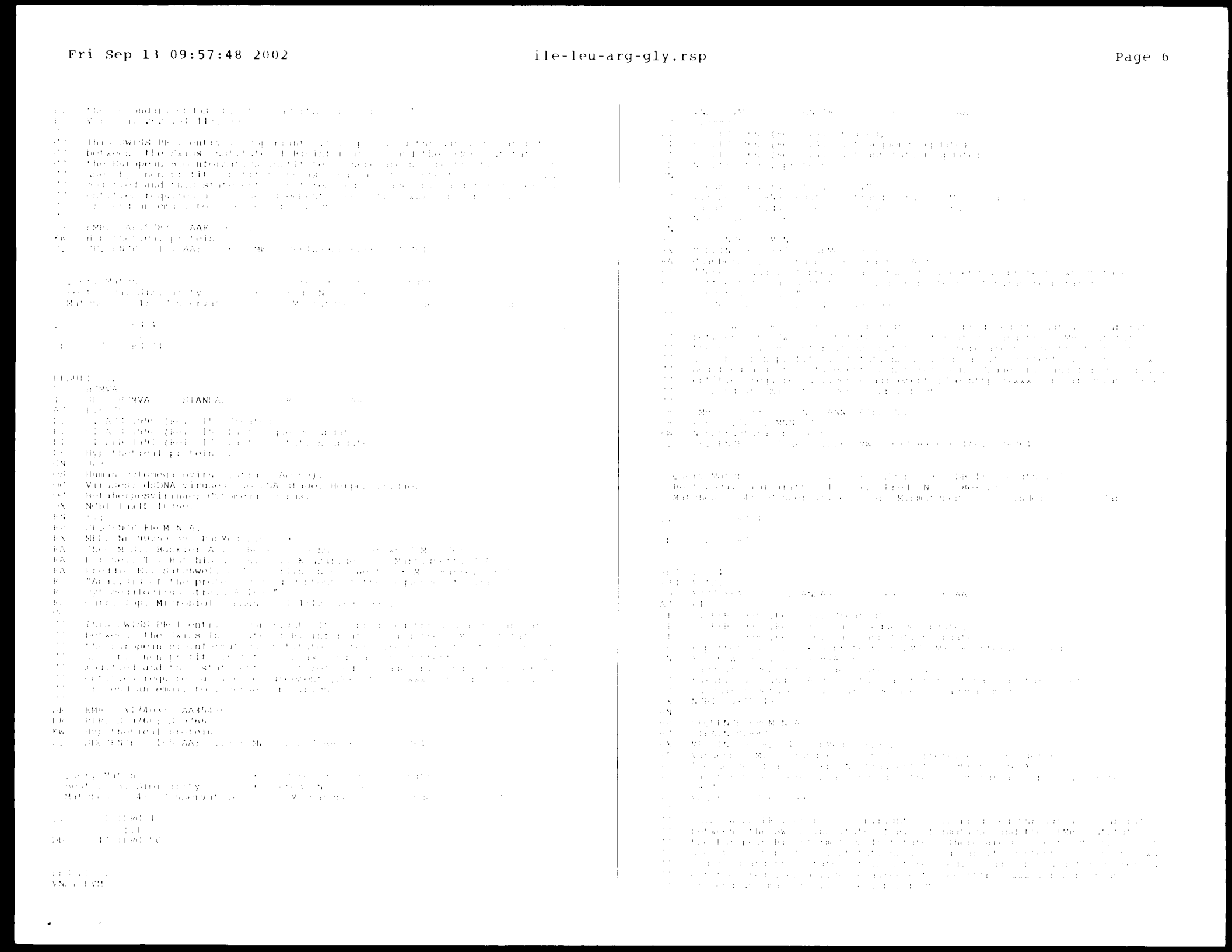
Qy 1 ILRG 4
||||
Db 19 ILRG 22

RESULT 2
BDS1_ANESU
ID BDS1_ANESU STANDARD; PRT; [illegible] AA.
AC P11494;
DT 01-OCT-1989 (Rel. 11, Created)
DT 01-OCT-1989 (Rel. 11, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Antihypertensive and antiviral protein BDS1.
OS Anemonia sulcata (Snakelocks sea anemone).
OC Eukaryota; Metazoa; Cnidaria; Anthozoa; [illegible]
OC Nynantheae; Actiniidae; Anemonia.

```
RA   Robison K.;
RT   Submitted (DEC 1994) to EMBL/GenBank/DDBJ databases.
RN   [6]
RP   CHARACTERIZATION, AND GENE NAME.
RX   MEDLINE=95362656; PubMed=7635817;
RA   Thony-Meyer L., Fischer F., Kunzler P., Ritz D., Hennecke H.;
RT   "Escherichia coli genes required for cytochrome c maturation.";
RL   J. Bacteriol. 177:4321-4326(1995).
CC   -!- FUNCTION: REQUIRED FOR THE EXPORT OF HEMES H AND B FROM THE CYTOPLASM FOR THE
CC       BIOGENESIS OF C-TYPE CYTOCHROMES (POTENTIAL).
CC   -!- SIMILARITY: BELONGS TO THE CCMB/DSBE FAMILY.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; U00008; AAA16390.1; -.
DR   EMBL; AE000309; AAC75258.1; -.
DR   EMBL; AE005452; AAG57333.1; -.
DR   EMBL; AP002560; BAB36510.1; -.
DR   EcoGene; EG12169; ccmD.
KW   Cytochrome c-type biogenesis; Transport; Transmembrane;
KW   Inner membrane; Complete proteome.
FT   TRANSMEM     18     37       POTENTIAL.
SQ   SEQUENCE   69 AA;  7745 MW;  BEC1D341EEADA018 CRC64;


  Query Match             100.0%;  Score 19;  DB 1;  Length 69;
  Best Local Similarity   100.0%;  Pred. No. 1.1e+02;
  Matches    4;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 ILRG 4
           ||||
Db      46 ILRG 49


RESULT   5
YF67_THEMA
ID   YF67_THEMA     STANDARD;      PRT;    75 AA.
AC   Q9X1Q3;
DT   30-MAY-2000 (Rel. 39, Created)
DT   30-MAY-2000 (Rel. 39, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Hypothetical protein TM1567.
GN   TM1567.
OS   Thermotoga maritima.
OC   Bacteria; Thermotogales; Thermotogaceae.
OX   NCBI_TaxID=2336;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=MSB8 / DSM 3109;
RX   MEDLINE=99287316; PubMed=10360571;
RA   Nelson K.E., Clayton R.A., Gill S.R., Gwinn M.L., Dodson R.J.,
RA   Haft D.H., Hickey E.K., Peterson J.D., Nelson W.C., Ketchum K.A.,
RA   McDonald L., Utterback T.R., Malek J.A., Linher K.D., Garrett M.M.,
RA   Stewart A.M., Cotton M.D., Pratt M.S., Phillips C.A., Richardson D.,
RA   Heidelberg J., Sutton G.G., Fleischmann R.D., Eisen J.A., White O.,
RA   Salzberg S.L., Smith H.O., Venter J.C., Fraser C.M.;
RT   "Evidence for lateral gene transfer between Archaea and Bacteria from
RT   genome sequence of Thermotoga maritima.";
RL   Nature 399:323-329(1999).
CC   -!- SIMILARITY: BELONGS TO THE UPF0109 FAMILY.
CC   -!- SIMILARITY: CONTAINS 1 KH DOMAIN.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; AE001802; AAL98044.1; -.
DR   TIGR; TM1567; -.
DR   InterPro; IPR003607; KH.
DR   InterPro; IPR004088; KH_TYPE_1.
DR   Pfam; PF00013; KH-domain; 1.
DR   PROSITE; PS50084; KH_TYPE_1; 1.
KW   Hypothetical protein; RNA-binding; Complete proteome.
FT   DOMAIN       27     91       KH.
SQ   SEQUENCE   75 AA;  8394 MW;  7AB2E1D6DE5...  CRC64;


  Query Match             100.0%;  Score 19;  DB 1;  Length 75;
  Best Local Similarity   100.0%;  Pred. No. 1.1e+02;
  Matches    4;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 ILRG 4
           ||||
Db       8 ILRG 11


RESULT   6
RUXG_HUMAN
ID   RUXG_HUMAN     STANDARD;      PRT;    72 AA.
AC   Q15357;
DT   15-JUL-1999 (Rel. 38, Created)
DT   15-JUL-1999 (Rel. 38, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Small nuclear ribonucleoprotein G (snRNP-G) (Sm protein G) (Sm-G)
DE   (SmG).
GN   SNRPG OR PBSCG.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=95262647; PubMed=7744013;
RA   Hermann H., Fabrizio P., Raker V.A., Foulaki K., Hartmuth C., Reuter R.,
RA   Luehrmann R.;
RT   "snRNP Sm proteins share two evolutionarily conserved sequence motifs
RT   which are involved in Sm protein-protein interactions.";
RL   EMBO J. 14:2076-2088(1995).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Placenta;
RA   Strausberg R.;
RL   Submitted (NOV 2000) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: THIS PROTEIN IS ASSOCIATED WITH SMALL NUCLEAR RNA'S U1, U2,
CC       U5.
CC   -!- SUBCELLULAR LOCATION: Nuclear.
CC   -!- SIMILARITY: BELONGS TO THE SNRNP SM PROTEIN FAMILY.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; X85973; CAA59689.1; -.
DR   EMBL; BC000070; AAH00070.1; -.
DR   MIM; 603542; -.
DR   InterPro; IPR001163; snRNP_Sm.
DR   Pfam; PF01423; Sm; 1.
KW   Nuclear protein; Ribonucleoprotein; mRNA processing; mRNA splicing;
KW   RNA-binding.
SQ   SEQUENCE   72 AA;  8196 MW;  62330ZBF6B5F...  CRC64;
```



```
          |||
Db      70 ILRG 73


RESULT   9
YC76_ARCFU
ID   YC76_ARCFU     STANDARD;      PRT;    94 AA.
AC   O28992;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Hypothetical protein AF1276 precursor.
GN   AF1276.
OS   Archaeoglobus fulgidus.
OC   Archaea; Euryarchaeota; Archaeoglobales; Archaeoglobaceae;
OC   Archaeoglobus.
OX   NCBI_TaxID=2234;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=VC 16 / DSM 4304 / ATCC 49558;
RX   MEDLINE=98049348; PubMed=9389475;
RA   Klenk H.-P., Clayton R.A., Tomb J.-F., White O., Nelson K.E.,
RA   Ketchum K.A., Dodson R.J., Gwinn M., Hickey E.K., Peterson J.D.,
RA   Richardson D.L., Kerlavage A.R., Graham D.E., Kyrpides N.C.,
RA   Fleischmann R.D., Quackenbush J., Lee N.H., Sutton G.G., Gill S.,
RA   Kirkness E.F., Dougherty B.A., McKenney K., Adams M.D., Loftus B.,
RA   Peterson S., Reich C.I., McNeil L.K., Badger J.H., Glodek A., Zhou L.,
RA   Overbeek R., Gocayne J.D., Weidman J.F., McDonald L., Utterback T.,
RA   Cotton M.D., Spriggs T., Artiach P., Kaine B.P., Sykes S.M.,
RA   Sadow P.W., D'Andrea K.P., Bowman C., Fujii C., Garland S.A.,
RA   Mason T.M., Olsen G.J., Fraser C.M., Smith H.O., Woese C.R.,
RA   Venter J.C.;
RT   "The complete genome sequence of the hyperthermophilic, sulphate-
RT   reducing archaeon Archaeoglobus fulgidus.";
RL   Nature 390:364-370(1997).
CC   -----------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   -----------------------------------------------------------------
DR   EMBL; AE001016; AAB89977.1;  -.
DR   TIGR; AF1276;  -.
KW   Hypothetical protein; Signal; Complete proteome.
FT   SIGNAL       1     25       POTENTIAL.
FT   CHAIN       26     94       HYPOTHETICAL 1P  3 N AF1276.
SQ   SEQUENCE   94 AA;  10151 MW;  A40F0DFD6957C4..  CRC64;


  Query Match             100.0%;  Score 19;  D  0  Length 94;
  Best Local Similarity   100.0%;  Pred. No. 1.6e-..;
  Matches    4;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 ILRG 4
           ||||
Db      40 ILRG 43


RESULT  10
TI12_SINAL
ID   TI12_SINAL     STANDARD;      PRT;    99 AA.
AC   P26780;
DT   01-AUG-1992 (Rel. 23, Created)
DT   01-OCT-1996 (Rel. 34, Last sequence update)
DT   15-DEC-1998 (Rel. 37, Last annotation update)
DE   Trypsin inhibitor 2 precursor (MII-2).
OS   Sinapis alba (White mustard) (Brassica hirta).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC   eurosids II; Brassicales; Brassicaceae; Sinapis.
OX   NCBI_TaxID=3728;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=CV. ALBAMOS; TISSUE=Seed;
RX   MEDLINE=96266796; PubMed=8706666;
RA   Ceci L.R., Sposito M., De Virgilio M., Di Fonzo N., Volpicella M.;
RT   "The gene coding for the mustard trypsin inhibitor 2 is expressed in
RT   and wound inducible.";
RL   FEBS Lett. 364:179-181(1995).
RN   [2]
RP   SEQUENCE OF 31-93.
RC   STRAIN=CV. ALBAMOS; TISSUE=Seed;
RX   MEDLINE=93086687; PubMed=1451776;
RA   Menegatti E., Tedeschi G., Ronchi S., Bossi S., Simoncini L.,
RA   Thomas R.M., Bollini R., Palmieri S.;
RT   "Purification, inhibitory properties and primary structure of a
RT   serine proteinase inhibitor from white mustard (Sinapis alba L.)
RT   seed.";
RL   FEBS Lett. 330:10-14(1993).
CC   -!- FUNCTION: A TRYPSIN INHIBITOR WITH ANTIFUNGAL PROPERTIES.
CC   -!- SUBCELLULAR LOCATION: ON A 1:1 MOLAR BASIS.
CC   -!- SIMILARITY: BELONG TO THE KUNITZ/MII-2 PROTEASE INHIBITOR FAMILY.
CC   -----------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   -----------------------------------------------------------------
DR   EMBL; X84208; CAA58941.1;  -.
DR   PIR; S21120; S21120.
DR   InterPro; IPR003614; Kunitz.
DR   SMART; SM00505; Kunitz; 1.
KW   Serine protease inhibitor; Signal.
FT   SIGNAL       1     30
FT   CHAIN       31     93       TRYPSIN INHIBITOR 2.
FT   PROPEP      94     99
FT   DISULFID    34     87       BY SIMILARITY.
FT   DISULFID    47     73       BY SIMILARITY.
FT   DISULFID    53     77       BY SIMILARITY.
FT   DISULFID    60     86       BY SIMILARITY.
FT   ACT_SITE    50     51       REACTIVE BOND (BY SIMILARITY).
SQ   SEQUENCE   99 AA;  [illegible] MW;  5E96C788D9F4 CRC64;


  Query Match             [illegible]  Score [illegible]  Length 94;
  Best Local Similarity   [illegible]  Pred. No. [illegible]
  Matches    4;  Conservative    [illegible]  Mismatches    [illegible]  Indels    [illegible]  Gaps    [illegible]

Qy       1 ILRG 4
           ||||
Db      94 ILRG 97


RESULT  11
VP15_BPAPS
ID   VP15_BPAPS     STANDARD;      PRT;    [illegible]
AC   Q91113;
DT   [illegible] (Rel. [illegible], Created)
DT   [illegible] (Rel. [illegible], Last sequence update)
DT   [illegible] (Rel. [illegible], Last annotation update)
DE   Putative protein P15.
GN   15.
OS   Bacteriophage Aaf3 [illegible].
OC   Viruses; dsDNA viruses, [illegible]; Caudovirales; Podoviridae;
OC   Picovirinae; [illegible].
OX   NCBI_TaxID=[illegible];
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=99423868; PubMed=[illegible];
RA   van der Wilk F., Dortmans A.M., Verbeet [illegible]
RT   "Isolation and characterization of AFS1 [illegible] bacteriophage
```



```
CC   ----------------------------------------------------------------------
DR   EMBL; Z34288; CAA84156.1; -.
DR   EMBL; Z49342; CAA89457.1; -.
DR   PIR; S47124; S47124.
DR   SGD; S0003603; YJL067W.
KW   Hypothetical protein; Transmembrane.
FT   TRANSMEM      43      65       POTENTIAL.
FT   TRANSMEM      72      92       POTENTIAL.
SQ   SEQUENCE   116 AA;  13214 MW;  [illegible] CRC64;


  Query Match             100.0%;  Score 19;  DB 1;  Length 116;
  Best Local Similarity   100.0%;  Pred. No. 1.9e+02;
  Matches    4;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 ILRG 4
           ||||
Db      59 ILRG 62


RESULT  15
RL24_METTH
ID   RL24_METTH     STANDARD;      PRT;   117 AA.
AC   O26122;
DT   15-JUL-1998 (Rel. 36, Created)
DT   15-JUL-1998 (Rel. 36, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   50S ribosomal protein L24P.
GN   RPL24P OR MTH14.
OS   Methanobacterium thermoautotrophicum.
OC   Archaea; Euryarchaeota; Methanobacteriales; Methanobacteriaceae;
OC   Methanothermobacter.
OX   NCBI_TaxID=145262;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN DELTA H;
RX   MEDLINE=98037514; PubMed=9371463;
RA   Smith D.R., Doucette-Stamm L.A., Deloughery C., Lee H., Dubois J.,
RA   Aldredge T., Bashirzadeh R., Blakely D., Cook R., Gilbert K.,
RA   Harrison D., Hoang L., Keagle P., Lumm W., Pothier B., Qiu D.,
RA   Spadafora R., Vicare R., Wang Y., Wierzbowski J., Gibson R.,
RA   Jiwani N., Caruso A., Bush D., Safer H., Patwell D., Prabhakar S.,
RA   McDougall S., Shimer G., Goyal A., Pietrovski S., Church G.M.,
RA   Daniels C.J., Mao J.-I., Rice P., Nolling J., Reeve J.N.;
RT   "Complete genome sequence of Methanobacterium thermoautotrophicum
RT   deltaH: functional analysis and comparative genomics.";
RL   J. Bacteriol. 179:7135-7155(1997).
CC   -!- SIMILARITY: BELONGS TO THE L24P FAMILY OF RIBOSOMAL PROTEINS.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; AE000795; AAB84515.1; -.
DR   InterPro; IPR003256; Ribosomal_L24.
DR   InterPro; IPR000402; Ribosomal_L24_NusG.
DR   Pfam; PF00467; Ribosomal_L24; 1.
DR   ProDom; PD001677; Ribosomal_L24; 1.
DR   PROSITE; PS01108; RIBOSOMAL_L24; FALSE_NEG.
KW   Ribosomal protein; Complete proteome.
SQ   SEQUENCE   117 AA;  13886 MW;  [illegible] CRC64;


  Query Match             100.0%;  Score 19;  DB 1;  Length 117;
  Best Local Similarity   100.0%;  Pred. No. 1.9e+02;
  Matches    4;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 ILRG 4
           ||||
Db      59 ILRG 62


RESULT  16
RL24_METJA
ID   RL24_METJA     STANDARD;      PRT;   117 AA.
AC   Q58124;
DT   15-JUL-1998 (Rel. 36, Created)
DT   15-JUL-1998 (Rel. 36, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   50S ribosomal protein L24P.
GN   RPL24P OR MJ1464.
OS   Methanococcus jannaschii.
OC   Archaea; Euryarchaeota; Methanococcales; Methanococcaceae;
OC   Methanococcus.
OX   NCBI_TaxID=2190;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN DSM 2661;
RX   MEDLINE=96326488; PubMed=8687420;
RA   Bult C.J., White O., Olsen G.J., Zhou L., Fleischmann R.D., Sutton G.G.,
RA   Blake J.A., FitzGerald L.M., Clayton R.A., Gocayne J.D., Kerlavage A.R.,
RA   Dougherty B.A., Tomb J.-F., Adams M.D., Reich C.I., Overbeek R.,
RA   Kirkness E.F., Weinstock K.G., Merrick J.M., Glodek A., Scott J.L.,
RA   Geoghagen N.S.M., Weidman J.F., Fuhrmann J.L., Nguyen D., Utterback T.R.,
RA   Kelley J.M., Peterson J.D., Sadow P.W., Hanna M.C., Cotton M.D.,
RA   Roberts K.M., Hurst M.A., Kaine B.P., Borodovsky M., Klenk H.-P.,
RA   Fraser C.M., Smith H.O., Woese C.R., Venter J.C.;
RT   "Complete genome sequence of the methanogenic archaeon, Methanococcus
RT   jannaschii.";
RL   Science 273:1058-1073(1996).
CC   -!- SIMILARITY: BELONGS TO THE L24P FAMILY OF RIBOSOMAL PROTEINS.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; L77117; AAB98795.1; -.
DR   InterPro; IPR003256; Ribosomal_L24.
DR   InterPro; IPR000402; Ribosomal_L24_NusG.
DR   Pfam; PF00467; Ribosomal_L24; 1.
DR   ProDom; PD001677; Ribosomal_L24; 1.
DR   PROSITE; PS01108; RIBOSOMAL_L24; FALSE_NEG.
KW   Ribosomal protein; Complete proteome.
SQ   SEQUENCE   117 AA;  13886 MW;  [illegible] CRC64;

Search completed: September [illegible]
Total time: [illegible]
```

GenCore version 4.5

OM protein - protein search, using sw model

Run on: September 13, 2002, 09:24:10 ; Search Time 15.13 Seconds
(without alignments)
25.367 Million cell updates/sec

Title: ILE-LEU-ARG-GLY
Perfect score: 19
Sequence: 1 ILRG 4

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 282198 seqs, 96090224 residues

Total number of hits satisfying chosen parameters 282198

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_71:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Prod. N: is the number of results predicted at or above this
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	19	100.0	34	2	A44038	leader peptide TnaC
2	19	100.0	39	2	G85837	hypothetical prote
3	19	100.0	43	2	S29763	aminopeptidase N -
4	19	100.0	43	2	A33041	antiviral protein
5	19	100.0	50	2	T31019	hypothetical prote
6	19	100.0	51	2	S55449	GAM1 protein - Pla
7	19	100.0	52	2	T08497	hypothetical prote
8	19	100.0	52	2	AE0410	hypothetical prote
9	19	100.0	53	2	B83853	hypothetical prote
10	19	100.0	57	2	S55452	GAM1 protein - Pla
11	19	100.0	58	2	T34120	hypothetical prote
12	19	100.0	61	2	F81510	hypothetical prote
13	19	100.0	62	2	F82261	hypothetical prote
14	19	100.0	68	2	S55451	GAM1 protein - Pla
15	19	100.0	68	2	S55455	GAM1 protein - Pla
16	19	100.0	69	1	D64989	heme exporter prot
17	19	100.0	69	2	G91014	heme exporter prot
18	19	100.0	69	2	A85859	heme exporter prot
19	19	100.0	70	2	AF0960	heme exporter prot
20	19	100.0	70	2	A10787	heme exporter prot
21	19	100.0	75	2	H72236	conserved hypothet
22	19	100.0	75	2	E97542	hypothetical prote
23	19	100.0	75	2	AF2761	hypothetical prote
24	19	100.0	76	2	S55954	Sm protein G - Ba
25	19	100.0	77	2	I51529	integrin alpha V-1
26	19	100.0	77	2	S48322	SNF2 protein - yea
27	19	100.0	77	2	AD0343	hypothetical prote
28	19	100.0	77	2	AI2248	hypothetical prote
29	19	100.0	79	2	S55456	GAM1 protein - Pla
30	19	100.0	85	2	A10787	heme exporter prot
31	19	100.0	85	2	AC2418	hypothetical pr
32	19	100.0	86	2	JC1160	DNA polymerase
33	19	100.0	87	2	I04128	probable DNA dir
34	19	100.0	89	2	AG2240	hypothetical pr
35	19	100.0	90	2	G00413	probable kinase
36	19	100.0	90	2	S09252	hypothetical pr
37	19	100.0	91	2	I41169	DNA primase homo
38	19	100.0	92	2	H75359	hypothetical pr
39	19	100.0	96	2	C94459	hypothetical pr
40	19	100.0	97	2	I98155	hypothetical protein
41	19	100.0	97	2	AC2927	phosphotyrosyl p
42	19	100.0	99	2	S85661	transcription
43	19	100.0	104	2	H70541	hypothetical pr
44	19	100.0	105	2	S09766	hypothetical protein
45	19	100.0	135	2	I71514	hypothetical pr

ALIGNMENTS

RESULT 1
A44038
leader peptide TnaC - Proteus vulgaris
C;Species: Proteus vulgaris
C;Date: 10-Jun-1993 #sequence_revision 18-Nov-1994 #text_change 24-Sep-1999
C;Accession: A44038
R;Kamath, A.V.; Yanofsky, C.
J. Biol. Chem. 267, 19978-19985, 1992
A;Title: Characterization of the tryptophanase operon leader peptide from Proteus vulgaris and its
A;Reference number: A44038; MUID:93015858
A;Accession: A44038
A;Status: preliminary; not compared with conceptual translati
A;Molecule type: nucleic acid
A;Residues: 1-34 <KAM>
A;Cross-references: GB:M77779; NID:g153919; PIDN:AAA25605.1;
A;Note: Sequence extracted from NCBI backbone (NCBI
C;Superfamily: unassigned leader peptides

Query Match 100.0%; Score 19; DB 2; Length 34;
Best Local Similarity 100.0%; Pred. No. 1.7e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ILRG 4
||||
Db 10 ILRG 13

RESULT 2
G85837
hypothetical protein Z3427 [imported] - Escherichia coli (strain O157:H7, substrain E
C;Species: Escherichia coli
C;Date: 19-Feb-2001 #sequence_revision 19-Feb-2001 #text_change 14-Sep-2001
C;Accession: G85837
R;Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew,
Blattner, F.R.; Barrell, B.G.; Etheridge, N.W.; Umenhoffer, K.; Pati, A.; Ayers, K.; Brown,
Nature 409, 529-533, 2001
A;Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7
A;Reference number: A85486; MUID:21074695; PMID:11206551
A,Accession: G85837
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-39 <STO>
A,Cross-references: GB:AE005174; NID:g12513412; PIDN:AAG57218.1; GSPDB:GN00135; GenBa
A,Experimental source: strain O157:H7, substrain EDL933
C;Genetics:
A,Gene: 23270

Query Match 100.0%; Score 19; DB 2; Length 39;
Best Local Similarity 100.0%; Pred. No. 1.7e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
C;Date: 02 Nov 2001 #sequence_revision 02 Nov 2001 #text_change 02 Nov 2001
C;Accession: AE0410
R;Parkhill, J.; Wren, B.W.; Thomson, N.R.; Titball, R.W.; Holden, M.T.G.; Prentice, M.B.
deno-Tarraga, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Davis, P.; Dougan, G.;
il, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrell,
Nature 413, 523-527, 2001
A;Title: Genome sequence of Yersinia pestis, the causative agent of plague.
A;Reference number: AB0001; MUID:21470413; PMID:[illegible]
A;Accession: AE0410
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-52 <KUR>
A;Cross-references: GB:AL590842; PIDN:CAC92609.1; [illegible]; GSPDB:GN00[illegible]
C;Genetics:
A;Gene: YPO3379

  Query Match             100.0%;  Score 19;  DB [illegible];  Length 52;
  Best Local Similarity   100.0%;  Pred. No. [illegible];
  Matches    4;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 ILRG 4
              ||||
Db          9 ILRG 12

RESULT      9
B83853
hypothetical protein BH1626 [imported] - Bacillus halodurans (strain C-125)
C;Species: Bacillus halodurans
C;Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 15-Jun-2001
C;Accession: B83853
R;Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fuji, F.; [illegible]
Nucleic Acids Res. 28, 4317-4331, 2000
A;Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans and
A;Reference number: A83650; MUID:20512582; PMID:[illegible]
A;Accession: B83853
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-53 &lt;STO&gt;
A;Cross-references: GB:AP001512; GB:BA000004; NID:[illegible]; PIDN:BAB0[illegible]; GSPDB:GN00
A;Experimental source: strain C-125
C;Genetics:
A;Gene: BH1626

  Query Match             100.0%;  Score 19;  DB [illegible];  Length 53;
  Best Local Similarity   100.0%;  Pred. No. [illegible];
  Matches    4;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 ILRG 4
              ||||
Db         20 ILRG 23

RESULT     10
S55452
GAM1 protein - Plasmodium vivax (isolate K1[illegible]) (fragment)
C;Species: Plasmodium vivax
C;Date: 06-Sep-1996 #sequence_revision 13-Mar-1997 #text_change 17-Mar-1999
C;Accession: S55452
R;Snewin, V.A.; Khouri, E.; Wattavidanage, [illegible]
Mol. Biochem. Parasitol. 71, 135-138, 1995
A;Title: A new polymorphic marker for PCR typing [illegible] Plasmodium vivax parasites.
A;Reference number: S55449; MUID:95356790
A;Accession: S55452
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-57 <SNE>
A;Cross-references: EMBL:X84673

  Query Match             100.0%;  Score 19;  DB [illegible];  Length 57;
  Best Local Similarity   100.0%;  Pred. No. [illegible];
  Matches    4;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 ILRG 4
              ||||
Db         39 ILRG 42

RESULT     11
T34120
hypothetical protein C26B2.5 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change [illegible]
C;Accession: T34120
R;Du, Z.; Gattung, S.
submitted to the EMBL Data Library, November [illegible]
A;Description: The sequence of C. elegans cosmid C26B2.
A;Reference number: Z21486
A;Accession: T34120
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-58 <DUZ>
A;Cross-references: EMBL:[illegible]; PIDN:AAC24[illegible]; [illegible]
A;Experimental source: strain Bristol N2; [illegible]
C;Genetics:
A;Gene: CESP:C26B2.5
A;Map position: 4
A;Introns: 20/3; 36/2

  Query Match             [illegible];  Score [illegible];  [illegible];  Length [illegible];
  Best Local Similarity   [illegible];  Pred. No. [illegible];
  Matches    4;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 ILRG 4
              ||||
Db          5 ILRG 8

RESULT     12
F81510
hypothetical protein CP1027 [imported] - Chlamydophila pneumoniae (strain AR39)
C;Species: Chlamydophila pneumoniae, Chlamydia pneumoniae
C;Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 11-May-2000
C;Accession: F81510
R;Read, T.D.; Brunham, R.C.; Shen, C.; Gill, [illegible]
, C.; Dodson, R.; Gwinn, M.; Nelson, W.; DeBoy, [illegible]
Nucleic Acids Res. 28, 1397-1406, 2000
A;Title: Genome sequences of Chlamydia trachomatis MoPn and Chlamydia pneumoniae AR39
A;Reference number: A81500; MUID:20154255
A;Accession: F81510
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-61 <REA>
A;Cross-references: GB:AE00[illegible]; GB:AE00[illegible]; [illegible]
A;Experimental source: strain AR39; [illegible]
C;Genetics:
A;Gene: CP1027

  Query Match             [illegible]
  Best Local Similarity   [illegible]
  Matches    4;  [illegible]

Qy          1 ILRG 4
              ||||
Db         4[illegible] ILRG 4[illegible]

RESULT     13
F82361
hypothetical protein V[illegible] [imported] - V[illegible] (strain N16961 [illegible])
C;Species: Vibrio cholerae
```

GenCore version 4.5

OM protein - protein search, using sw model

Run on: September 13, 2002, 09:19:40 ; Search time 29.91 Seconds
(without alignments)
4.504 Million cell updates/sec

Title: ILE-LEU-ARG-GLY
Perfect score: 19
Sequence: 1 ILRG 4

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 747574 seqs, 111073796 residues

Total number of hits satisfying chosen parameters: 747574

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_032802:*
1: /SIDS1/gcgdata/bl2d/geneseqp-embl/AA1980.DAT:*
2: /SIDS1/gcgdata/bl2d/geneseqp-embl/AA1981.DAT:*
3: /SIDS1/gcgdata/bl2d/geneseqp-embl/AA1982.DAT:*
4: /SIDS1/gcgdata/bl2d/geneseqp-embl/AA1983.DAT:*
5: /SIDS1/gcgdata/bl2d/geneseqp-embl/AA1984.DAT:*
6: /SIDS1/gcgdata/bl2d/geneseqp-embl/AA1985.DAT:*
7: /SIDS1/gcgdata/bl2d/geneseqp-embl/AA1986.DAT:*
8: /SIDS1/gcgdata/bl2d/geneseqp-embl/AA1987.DAT:*
9: /SIDS1/gcgdata/bl2d/geneseqp-embl/AA1988.DAT:*
10: /SIDS1/gcgdata/bl2d/geneseqp-embl/AA1989.DAT:*
11: /SIDS1/gcgdata/bl2d/geneseqp-embl/AA1990.DAT:*
12: /SIDS1/gcgdata/bl2d/geneseqp-embl/AA1991.DAT:*
13: /SIDS1/gcgdata/bl2d/geneseqp-embl/AA1992.DAT:*
14: /SIDS1/gcgdata/bl2d/geneseqp-embl/AA1993.DAT:*
15: /SIDS1/gcgdata/bl2d/geneseqp-embl/AA1994.DAT:*
16: /SIDS1/gcgdata/bl2d/geneseqp-embl/AA1995.DAT:*
17: /SIDS1/gcgdata/bl2d/geneseqp-embl/AA1996.DAT:*
18: /SIDS1/gcgdata/bl2d/geneseqp-embl/AA1997.DAT:*
19: /SIDS1/gcgdata/bl2d/geneseqp-embl/AA1998.DAT:*
20: /SIDS1/gcgdata/bl2d/geneseqp-embl/AA1999.DAT:*
21: /SIDS1/gcgdata/bl2d/geneseqp-embl/AA2000.DAT:*
22: /SIDS1/gcgdata/bl2d/geneseqp-embl/AA2001.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	19	100.0	5	21	AAY87478	Antheraea yamamai
2	19	100.0	5	22	AAB58532	Antheraea yamamai
3	19	100.0	9	22	AAB87527	Substituted phen
4	19	100.0	9	15	AAY38009	Influenza virus-d
5	19	100.0	9	15	AAR73769	Antigen fragment 8
6	19	100.0	9	15	AAR73771	Antigen fragment 8
7	19	100.0	9	16	AAR70066	Control peptide 92
8	19	100.0	9	16	AAR70075	Hepatitis B virus
9	19	100.0	9	18	AAW39550	HBV pol peptide (p
10	19	100.0	9	18	AAW10386	Antiphospholipid a
11	19	100.0	9	19	AAW67106	Antibody ACA-[illegible]
12	19	100.0	9	19	AAW64641	Sept. de Ile-A-Leu
13	19	100.0	9	19	AAW61075	Amino acid sequenc
14	19	100.0	9	19	AAW64500	Synthetic polypep
15	19	100.0	9	20	AAY45579	Immunogenic peptid
16	19	100.0	9	20	AAY46005	Immunogenic peptid
17	19	100.0	9	20	AAY46415	Immunogenic peptid
18	19	100.0	9	20	AAY46450	Immunogenic peptid
19	19	100.0	9	20	AAY10554	[illegible] epitope/Mn
20	19	100.0	9	20	AAY10576	HLA Class I mol I
21	19	100.0	9	21	AAY96354	HLA A1 binding pep
22	19	100.0	9	21	AAY44423	Inf enza NP (A/S
23	19	100.0	9	22	AAB11844	Inf enza A virus
24	19	100.0	9	22	AAM06949	HLA A3 predicted i
25	19	100.0	10	15	AAY38018	Inf enza virus de
26	19	100.0	10	15	AAR61527	Sept. de Ile-Thr-Met (I
27	19	100.0	10	16	AAR70046	Hepatitis B virus
28	19	100.0	10	20	AAY15588	Immunogenic peptid
29	19	100.0	10	22	AAG86923	Saccharomyces cere
30	19	100.0	10	22	AAC89927	Saccharomyces cere
31	19	100.0	10	22	AAG87790	Saccharomyces cere
32	19	100.0	10	22	AAG87791	Saccharomyces cere
33	19	100.0	12	16	AAR66972	Human IL-4 mutant
34	19	100.0	14	19	AAW45884	Sept. de membrane i
35	19	100.0	7	21	AAY58861	Membrane bound-ty
36	19	100.0	14	19	AAW45896	Sept. de membrane i
37	19	100.0	14	21	AAY58866	Soluble [illegible] mut
38	19	100.0	14	22	AAM97491	Brain peptide #22
39	19	100.0	11	19	AAW05444	Bone [illegible] cyclo
40	19	100.0	15	21	AAY72992	Hepatitis B virus
41	19	100.0	15	22	AAB74919	[illegible] super pilin
42	19	100.0	5	11	AAW06735	Synthetic peptide #2
43	19	100.0	5	11	AAW09584	Synthetic peptide #2
44	19	100.0	11	11	AAU25954	Human thrombopoiet
45	19	100.0	15	15	AAW10217	Fibrinogen-beta-ami

ALIGNMENTS

RESULT 1
AAY87478
ID AAY87478 standard; protein; 5 AA.
XX
AC AAY87478;
XX
DT 03-OCT-2000 (first entry)
XX
DE Antheraea yamamai repressive factor, a dipeptide terminal amide.
XX
KW Antheraea yamamai repressive factor; Any IPS dispause terminal amide;
KW Japanese oak silkworm; control; cocoon; pupation.
XX
OS Antheraea yamamai var. Guerin-Meneville.
XX
FH Key Location/Qualifiers
FT Modified-site 5
FT /note= "C terminal amide"
XX
PN JP3023790-B1.
XX
PD 21-MAR-2000.
XX
PF 31-MAY-1999; 99JP-0152273.
XX
PR 31-MAY-1999; 99JP-0152273.
XX
PA (NOKO) NOKIN KOGYO KK.
XX
DR WPI; 2000-357161/31
XX
PT Gene Any-kP with insect dispause control [illegible] post
PT prevention .
XX

Query Match 100.0%; Score 19; DB [illegible] Length 8;
Best Local Similarity 100.0%; Pred. No. [illegible];
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ILRG 4
||||
Db 1 ilrg 4

RESULT 4
AAY38009
ID AAY38009 standard; Peptide; 9 AA.
XX
AC AAY38009;
XX
DT 29-SEP-1999 (first entry)
XX
DE Influenza virus derived HLA-binding peptide
XX
KW Immunogen; HLA; human leukocyte antigen; binding motif; antiviral;
KW MHC; major histocompatability complex; cytotoxic T cell; vaccine;
KW prostate cancer; lymphoma; hepatitis; AIDS; diagnostic; diagnosis
XX
OS Influenza virus.
XX
PN WO9403205-A1.
XX
PD 17-FEB-1994.
XX
PF 06-AUG-1993; 93WO-US07421.
XX
PR 05-MAR-1993; 93US-0027746.
PR 07-AUG-1992; 92US-0926666.
XX
PA (CYTE-) CYTEL CORP.
XX
PI Celis E, Grey HM, Kubo RT, Sette A;
XX
DR WPI; 1994-065403/08.
XX
PT Peptide which specifically binds selected MHC class I molecule used to
PT induce an immune response for treatment or prevention of viral
PT infection or cancer, or for diagnosis
XX
PS Disclosure; Page 105; 150pp; English.
XX
CC The sequence is a specific example of a group of new immunogenic
CC peptides having an HLA-A3.2, HLA-A1, HLA-A1 + HLA-A24.1 binding
CC motif. For example, the peptides having an A3.2 binding motif
CC each have 9-10 residues and contain, from N- to C-terminus to the
CC C-terminus, (a) a first conserved residue selected from L, M, I,
CC V, S, A, T, F, C, G, D and E and (b) a second conserved residue of
CC K, R, Y, H or F, where the first and second conserved residues are
CC separated by 6-7 residues. The peptides are capable of binding
CC selected MHC molecules and inducing an immune response. They can be
CC used to treat and/or prevent viral infection, cancer, e.g. prostate
CC cancer, lymphoma, hepatitis or AIDS. They can also be used to produce
CC antibodies for use as diagnostic or therapeutic agents. The peptides
CC can also be used as diagnostic agents.
XX
SQ Sequence 9 AA;

Query Match 100.0%; Score 19; DB [illegible] Length 9;
Best Local Similarity 100.0%; Pred. No. [illegible];
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ILRG 4
||||
Db 1 ilrg 4

RESULT 5
AAR73768
ID AAR73768 standard; peptide; 9 AA.
XX
AC AAR73768;
XX
DT 19-JUN-1995 (first entry)
XX
DE Antigen fragment 84, from HBV has binding affinity for [illegible]
XX
KW antigen; epitope; immunogenic target protein; HBsAg; HBV; HBVs; HBV;
KW HIV1; plasma specific antigen; hepatitis B virus; Epstein-Barr;
KW human immunodeficiency virus; human papilloma virus; pre-S; EBNA;
KW MAGE 1, melanoma antigen 1; core antigen; e-antigen;
KW pharmaceutical composition; in vivo; ex vivo; therapeutic;
KW diagnostic; MHC class I molecule; major histocompatibility; [illegible];
KW HLA-A2.1; 9mer; cancer; patient; human; [illegible]
KW algorithm prediction; MAGE; CMV; cytomegalovirus; HSV;
KW herpes simplex virus; influenza A; MI.
XX
OS Hepatitis B virus.
XX
PN WO9420127-A.
XX
PD 15-SEP-1994.
XX
PF 04-MAR-1994; 94WO-US02357.
XX
PR 05-MAR-1993; 93US-0027746.
PR 04-JUN-1993; 93US-0073205.
PR 29-NOV-1993; 93US-0159264.
XX
PA (CYTE-) CYTEL CORP.
XX
PI Grey HM, Kast WM, Sette A, Sidney
XX
DR WPI; 1994-302678/37.
XX
PT Immunogenic peptide(s) having an HLA-A... [illegible]
PT for treatment or prophylaxis of cancer [illegible]
PT autoimmune diseases.
XX
PS Disclosure; Page [illegible]
XX
CC AAR73685-876 are potential peptide binders [illegible]
CC motifs disclosed in the [illegible]
CC further motifs, each peptide with binding [illegible]
CC (binding affinity is expressed as an [illegible]
CC standard peptide (AAR7...) in assays [illegible]
CC polymerase has a binding [illegible]
CC can induce cytotoxic lymphocytes which [illegible]
CC They can be used for the treatment or prophylaxis of cancer, infection
CC cancer or lymphoma [illegible].
XX
SQ Sequence 9 AA;

Query Match [illegible]
Best Local Similarity [illegible]; Pred. No. [illegible];
Matches 4; Conservative [illegible]; Mismatches [illegible]; Indels [illegible]; Gaps [illegible];

Qy 1 ILRG 4
||||
Db 2 ilrg 5

RESULT 6
AAR73771
ID AAR73771 standard; peptide; 9 AA.
XX
AC AAR73771;
XX

PI New peptides inducing cytotoxic lymph. /c ... p...
PI virus - are regions of HB polymerase protein ... treating acute
PI and chronic infections
XX
PS Example 2; Page 39; 85pp; English.
XX
CC Peptides AAR70069-77 are peptides derived ... the Hepatitis B virus
CC (HBV) polymerase (HBpol) used in an assay ... detect HBpol epitopes that
CC have affinity for cytotoxic T cells (CTL) ... Class I haplotypes,
CC especially HLA-A2.1. The peptide presented ... is generated from
CC residues 772-780 of HBpol. The HBpol epitope peptides (see AAR70044-55)
CC can be used, prophylactically as vaccines, ... with, or conjugated
CC to, HBV helper epitopes (AAR70060-4). The ... ides can be used,
CC particularly ex vivo, to stimulate CTL ... cells which cells can be
CC reintroduced into patients who have chronic ... acute HBV infections or
CC are carriers, especially in treatments to prevent conversion from acute
CC to chronic infections.
XX
SQ Sequence ... 9 AA;

Query Match ... 100.0%; Score ...; ... ...; Length ...;
Best Local Similarity ... 100.0%; Pred. No. ...;
Matches ... 4; Conservative ... 0; Mismatches ... 0; Indels ... 0; Gaps ...

Qy 1 ILRG 4
||||
Db 2 ilrg 5

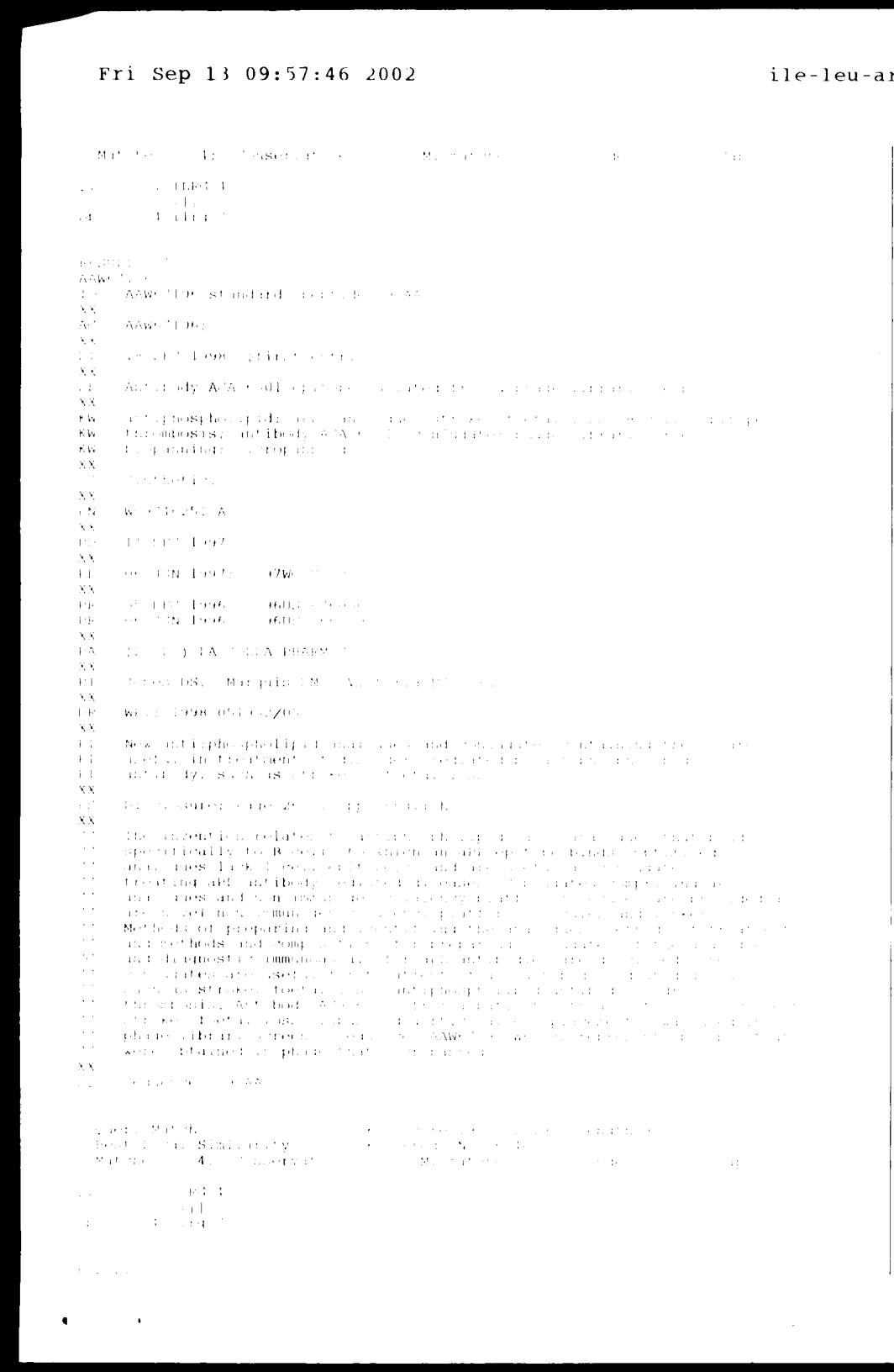
RESULT 9
AAW39550
ID AAW39550 standard; peptide; 9 AA.
XX
AC AAW39550;
XX
DT 11-JUN-1998 (first entry)
XX
DE HBV pol peptide (pos. 1344-1452) capable of binding to HLA-A*0201.
XX
KW T cell epitope; immune response; human leukocyte antigen; HLA class I;
KW vaccine; immunotherapy; major histocompatibility complex; MHC; B cell;
KW disease; anti-tumour; anti-viral.
XX
OS Hepatitis B Virus.
XX
PN WO9741440-A1.
XX
PD 06-NOV-1997.
XX
PF 28-APR-1997; 97WO-NL00229.
XX
PR 29-DEC-1996; 96EP-0203670.
PR 26-APR-1996; 96EP-0201145.
XX
PA (UYLE-) RIJKSUNIV LEIDEN.
PA (SCIS-) SCI SEED CAPITAL INVESTMENTS BV
XX
PI Kast WM, Meliet CJM, Ottringa R, Toes REM, Van Der Burg SH;
XX
DR WPI; 1997-549831/50.
XX
PT Method of selecting T cell peptide epitopes ... by measuring the
PT stability of HLA class I peptide complexes ... intact B cells
XX
PS Example 2; Page 63; 109pp; English.
XX
CC Peptides AAW39436-W39734 are used in a ... method for the selection of
CC immunogenic T-cell peptide epitopes present ... polypeptide antigens. The
CC method involves the identification of pept... sequences capable of
CC binding to an HLA (human leukocyte antigen ... class I molecule and
CC measuring the binding of this epitope pept... to the HLA class I
CC peptide. The stability of binding of the pe... ide and MHC (major
CC histocompatibility, ... ex) class I molecule is measured by intact human
CC B cells carrying the MHC molecule at the ... cell surfaces. The method can
CC be used to select peptide epitopes for ... ination against a
CC disease associated with the polypeptide, ...ssess of AIDS. The
CC peptide epitopes are especially T-cell pe... peptides with strong
CC anti-tumour and anti-viral immune responses. Peptides AAW39548-W39550 are
CC conserved hepatitis B derived peptides ... ... assay ... determine
CC immunogenicity and binding affinity to HLA-A*0201.
XX
SQ Sequence ... 9 AA;

Query Match ... ...; Score ...; ... ...; Length ...;
Best Local Similarity ... ...; Pred. No. ...;
Matches ... 4; Conservative ... 0; Mismatches ... 0; Indels ... 0; Gaps ...

Qy 1 ILRG 4
||||
Db 2 ilrg 5

RESULT 1
AAW14886
ID AAW14886 standard; peptide; 9 AA
XX
AC AAW14886;
XX
DT 21-AUG-1997 (first entry)
XX
DE Antiphospholipid ... peptide A^A ... shared ... ...
XX
KW aPL; antibody mediated path... system ... ...
KW antiphospholipid antibody syndrome; ... ...
XX
OS Synthetic.
XX
PN WO9640197-A1.
XX
PD 19-DEC-1996.
XX
PF 06-JUN-1996; ...
XX
PR 07-JUN-1995; ...
XX
PA (LJOL-) LA JOLLA INST.
XX
PI Jones DS, Margolis DM, Vanderslice P, ...
XX
DR WPI; 1997-071899/07.
XX
PT New antiphospholipid ... ... peptide prod... for the
PT diagnosis and treatment of aPL antibody ... pathologies such as
PT systemic lupus erythematosus
XX
PS Disclosure; Page ...; ... pp; English.
XX
CC The present sequence is an antiphospho... ...nal for the
CC 3E11 screened from A^A 650 phage library. Peptides are derived
CC from screening random peptide phage lib... ... codes
CC the sequences which best mimic the epitopes ... ... aPL
CC antibodies. A method has been produced ... ...
CC key epitopes recognised by aPL antibodies ... pathways in the
CC PAPS, APS and other aPL antibody mediated diseases. The peptides and
CC methods can be used in the diagnosis and ... ...
CC mediated pathologies such as systemic lupus erythematosus
CC antiphospholipid antibody syndrome ... ...
CC loss or thrombosis.
XX
SQ Sequence ... 9 AA;

Query Match ... ...; Score ...; ... ...; Length ...;
Best Local Similarity ... ...; Pred. No. ...;



PN WO9824460-A1.
XX
PD 04-JUN-1998.
XX
PF 25-NOV-1997; 97WO-GB03222.
XX
PR 25-NOV-1996; 96GB-0024456.
XX
PA (CITC) , IMPERIAL INNOVATION LTD.
XX
PI Brookes RH, Lalvani A;
XX
DR WPI; 1998-322871/28.
XX
PT Assay for peptide-specific T cells from cytotoxic T lymphocyte induced by
PT the peptide - used to monitor progress of infection or effect of
PT vaccination, for identifying T cell epitope(s) and for studying
PT infection
XX
PS Example 1; Page 9; 25pp; English.
XX
CC This is the nucleotide sequence of the Class I restricted influenza
CC epitopes recognised by freshly isolated CD8+ effector T cells. They
CC are used in the method of the invention, where the progress of
CC infection or the effect of a vaccination is monitored. The method is
CC specifically used to monitor progress of human immune deficiency virus
CC (HIV) (infection or the effect of vaccination) to identify
CC pathogen-derived epitopes targeted to CD4+ or CD8+ cells, or to study
CC hepatitis B or C, tuberculosis, malaria, HIV or influenza (e.g.
CC assessing whether a person who has recovered from an infection is
CC resistant to further infection).
XX
SQ Sequence 9 AA;

Query Match 100.0%; Score 19; DB 4; Length 9;
Best Local Similarity 100.0%; Pred. No. 6.4e-05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ILRG 4
||||
Db 1 ilrg 4

RESULT 14
AAW54500
ID AAW54500 standard; peptide; 9 AA.
XX
AC AAW54500;
XX
DT 01-SEP-1998 (first entry)
XX
DE Synthetic polypeptide HBV Pol 1344-1352.
XX
KW Carbohydrate-specific, cytolytic T cell, therapeutic, tumour,
KW major histocompatibility complex; vaccine; protective immune response;
KW pathogenic bacteria; virus; CTL response.
XX
OS Synthetic.
XX
PN WO9815287-A1.
XX
PD 16-APR-1998.
XX
PF 08-OCT-1997; 97WO-US18146.
XX
PR 08-OCT-1996; 96US-0028269.
XX
PA (LJOL-) LA JOLLA INST ALLERGY & IMMUNOLOGY.
XX
PI Grey H;
XX
DR WPI; 1998-216595/21.
XX
PT Composition for stimulating carbohydrate specific cytotoxic T
PT lymphocytes comprises synthetic peptide with attached cytolytic
PT hapten, designed to bind to class I molecule used in vaccines
PT against cancer or infectious disease
XX
PS Disclosure; Page [illegible]; [illegible]pp; English.
XX
CC The peptide AAW545[illegible] and AAW54[illegible] are the synthetic
CC polypeptides tested and used in the prototype of a synthetic
CC polypeptide, for stimulating a carbohydrate specific cytotoxic T cell
CC (CTL) response. The peptide should be at least 7 amino acids in
CC which are anchor residues. The polypeptide is the binding region
CC of a major histocompatibility complex class I molecule which is linked to an internal cytolytic hapten
CC molecule which is linked to the C-terminal residue of the peptide
CC extends beyond the anchor residues. This stimulates T cells which are able to
CC lyse specifically cells that express carbohydrate residues on their
CC surface. The polypeptides are used as vaccines to provide therapeutic
CC or protective immune responses, particularly against HIV, as for
CC against pathogenic bacteria and viruses (e.g. particularly those that cause
CC leprosy and tuberculosis, etc.).
XX
SQ Sequence 9 AA;

Query Match [illegible]; Score 19; DB [illegible]; Length
Best Local Similarity [illegible]; Pred. No. 6.4e-05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ILRG 4
||||
Db 2 ilrg 5

RESULT 15
AAY45579
ID AAY45579 standard; Peptide; 9 AA.
XX
AC AAY45579;
XX
DT 01-DEC-1998 (first entry)
XX
DE Immunogenic peptide derived a human leukocyte antigen bind molecule #19.
XX
KW Human leukocyte antigen; binding; immunogenic; polypeptide; MHC; HLA
KW immune response; T cell activation; macrophage, peptide
KW cytotoxic T lymphocyte; CTL tumour response; cell proliferation;
KW prostate cancer; hepatitis B; hepatitis C virus; retinal antigen;
KW vaccine; immunisation.
XX
OS Synthetic.
OS Homo sapiens.
XX
PN WO9945954-A1.
XX
PD 16-SEP-1999.
XX
PF 12-MAR-1999; 99WO-US05489.
XX
PR 13-MAR-1998; 98US-0077732.
XX
PA (EPIM) EPIMMUNE INC.
XX
PI Sette A, Kubo RT, Sidney J, Celis E, [illegible] BM, Southwood S;
XX
DR WPI; 1999-551214/46.
XX
PT New immunogenic peptides with HLA binding activity useful in treatment
PT and diagnosis of cancers and viral diseases
XX
PS Claim 1; Page 36; [illegible]pp; English.
XX
CC AAY45530 to AAY46214 represent specific human major histocompatibility